rla Central Library
PILANI (Jaipur State)
[redacted]
Class No :- 509 R
Book No :- B144S v.4
Accession No :- 45044

THE SOCIETY OF THE SIGMA XI

DEVOTED TO THE
ENCOURAGEMENT OF RESEARCH IN SCIENCE

National Lectureships
1943 and 1944

SCIENCE IN PROGRESS

FOURTH SERIES

☆ ☆ ☆ ☆

Science in Progress

By

WALTER R. MILES
DETLEV W. BRONK
SELIG HECHT
OTTO LOEWI
GEORGE D. BIRKHOFF
PETER DEBYE
HENRY EYRING
ISIDORE I. RABI
K. C. D. HICKMAN
C. A. ELVEHJEM
EDWIN J. COHN

Foreword by
LORANDE LOSS WOODRUFF

Edited by
GEORGE A. BAITSELL

FOURTH SERIES

☆ ☆ ☆ ☆

New Haven
YALE UNIVERSITY PRESS

LONDON · GEOFFREY CUMBERLEGE · OXFORD UNIVERSITY PRESS

Copyright, 1945, Yale University Press
Printed in the United States of America

First published, April, 1945
Second printing, July, 1945
Third printing, January, 1946
Fourth printing, January, 1947

All rights reserved. This book may not be reproduced in whole or in part, in any form (except by reviewers for the public press), without written permission from the publishers.

Preface

THE appearance of this volume of *Science in Progress,* the fourth of the series, is the response to a continued demand for the permanent publication of important contributions from ranking scientists in all fields, even though in some cases the articles necessarily include highly technical material. Several of the contributions published in the earlier volumes are already recognized as matured and essential portions of man's intellectual heritage; knowledge that will be of increasing value in the years ahead. It is believed that the material presented in this volume is of equal importance.

War research has taken the services of many scientists, particularly in the fields of chemistry and physics. Fortunately, some of the results of these investigations have now been released for publication. Such are the contributions in this volume of Drs. Miles, Bronk, and Cohn. All the articles printed here, with one exception, were prepared for the Sigma Xi National Lectureships and presented to widely distributed Chapter groups. Dr. Cohn's manuscript on "Blood and Blood Derivatives," which was the last to be received for inclusion in this book, was presented at Cleveland in September, 1944, as the Annual Sigma Xi Lecture.

Brief biographies of the contributors are given below:

George David Birkhoff, A.B., A.M. (Harvard), Ph.D. (University of Chicago). Formerly Perkins Professor of Mathematics at Harvard University. One of the world's greatest mathematicians, whose untimely death occurred just as this volume was going to press, Dr. Birkhoff possessed the rare faculty of putting the result of his profound mathematics in a form in his lectures and writings which interests the uninitiated and gives them a glimpse of the underlying principles. National Academy of Sciences.

Detlev W. Bronk, A.B. (Swarthmore), M.S., Ph.D. (Michigan). Professor of Biophysics, University of Pennsylvania; Director, Eldridge R. Johnson Foundation for Medical Physics; Director, Institute of Neurology, University of Pennsylvania; National Research Council; Coördinator of Research, Air Surgeon's Office. Dr. Bronk belongs to the small group of highly trained scientists capable of bridging the increasingly important fields of physics and biology. His services to aviation medicine during the present emergency are of the highest importance. National Academy of Sciences.

Edwin J. Cohn, B.S., Ph.D. (Chicago). Professor of Biochemistry, Harvard University, and Chairman of the Department of Physical Chemistry, Harvard Medical School. Dr. Cohn has made contributions in many fields of biochemistry, but during recent years has largely concentrated his attack upon the chemistry of proteins; particularly the chemical interactions between, and the separation of, the protein components in protein solutions. This field involves the proteins of blood plasma and here, in his war research, Dr. Cohn has been responsible for the direction of a large staff which has obtained results of major importance in the alleviation of human suffering. National Academy of Sciences.

Peter Debye, E.E. (Engineering School, Aachen), Ph.D. (Munich). Formerly Professor at Zurich, Utrecht, Goettingen, Leipzig, Berlin, and now at Cornell (Chairman, Department of Chemistry). In 1936 Dr. Debye received the Nobel Prize for his work in chemistry, and he was also elected as a foreign associate of the National Academy of Sciences. During the past few years a large portion of his time has been required for government research and consultation essential to the war program.

C. A. Elvehjem, B.S., Ph.D. (Wisconsin). Professor of Biochemistry at the University of Wisconsin. A ranking member of one of the most important and active departments of biochemistry in the world, Dr. Elvehjem has made many significant contributions to the field of nutrition, as attested by his recent election to the National Academy of Sciences.

Lately his investigations have been directed toward increasing our knowledge of the vitamin B complex.

Henry Eyring, B.S. (Arizona), Ph.D. (California). Professor of Physical Chemistry, Princeton University. A young man, Dr. Eyring is outstanding in one of the most difficult fields of physical chemistry, that of molecular behavior. Early in his career he was awarded the annual prize of the American Association for the Advancement of Science. Dr. Eyring is recognized as one of the leading members of the notable group in chemistry at Princeton. National Academy of Sciences.

Selig Hecht, B.S. (College of the City of New York), Ph.D. (Harvard). Professor of Biophysics, Columbia University. One of the foremost investigators in the field of vision, which requires knowledge of both physics and biology, Dr. Hecht has held lectureships at home and abroad and is a member of a number of learned societies, including the National Academy of Sciences. His investigations in the field of vision are directly related to the war, and he has therefore given a great deal of time during the past few years to the National Research Council and the Office of Scientific Research and Development.

K. C. D. Hickman, B.Sc., Ph.D. (Royal College of Science). Vice-President and Director of Research, Distillation Products Company, Rochester, N. Y. Coming from England for a position as research chemist for the Eastman Kodak Company (1925-38), Dr. Hickman soon made valuable contributions, particularly in the fields of photographic chemistry and vacuum distillation. The importance of this new field, based largely on Dr. Hickman's own discoveries, resulted in the formation of a new organization, the Distillation Products Company. Dr. Hickman has received wide recognition in scientific circles, and the results of his investigations, permitting rapid concentration of vitamins, have considerable commercial importance.

Otto Loewi, M. D. (University of Strassburg). Research Professor of Pharmacology, New York University, formerly

Professor of Pharmacology at the university at Gratz, Austria. Holder of internationally famous lectureships, member of many learned societies, and recipient of important honors in science, Dr. Loewi's investigations have covered many fields of human physiology. For his discovery of the chemical transmission of nerve impulse, as described in the present volume, he received the Nobel Prize in 1936.

Walter R. Miles, B.S. (Pacific College), A.B. (Earlham), M.A., Ph.D. (University of Iowa). Professor of Psychology, Yale University; President, Psychological Corporation; member, National Research Council, National Academy, and many other learned societies. Dr. Miles is an outstanding experimental psychologist, a notable inventor of psychological apparatus, and a frequent contributor to scientific journals. At present he is an essential adviser and investigator in studies of the psychological aspects of war aviation; the results will be of no less significance in peacetime.

Isidore I. Rabi, B.Chem. (Cornell), Ph.D. (Columbia). After completing work for his doctorate, Dr. Rabi studied abroad under several masters in the field of nuclear physics, then in 1929 returned to Columbia as a member of the staff and received his appointment as Professor of Physics in 1937. He has held several major lectureships, and is a member of the National Academy of Sciences. In 1936 he received the Sigma Xi Semi-Centennial Prize, in 1939 the prize of the American Association for the Advancement of Science, and in 1944 the Nobel Prize. He is now devoting all of his time to war research at the Massachusetts Institute of Technology.

The publication of *Science in Progress* is possible only by a highly coöperative effort involving authors, publisher, the *American Scientist,* and Sigma Xi Chapters and Clubs in every region of the country; to all these thanks are due and sincerely extended by the editor.

GEORGE A. BAITSELL
Executive Secretary, SIGMA XI

New Haven, Connecticut,
December, 1944.

Foreword

THE accumulation of knowledge during the millennia since man's emergence from a condition of complete dependence on his environment to the relatively masterful position that he now holds is, except in its broad sweep, beyond the compass of one mind, however talented. The polyhistor has been succeeded by the specialist.

But this obvious fact is, of course, a positive token of progress, even though it implies the loss of individual perspective as the *corpus scientiarum* necessarily becomes progressively divided into limited, workable fields of research. No one more fully realizes and deplores this restriction than the specialist himself who must largely depend for breadth of view on summaries made by his colleagues in fields other than his own. True, he may well envy, at least in his reveries, the investigators of a few centuries ago who, with limited background and facilities, could nevertheless pioneer in broad realms. However, with the scope of his endeavors less extensive, but more intensive, the modern specialist has full compensation in the satisfaction that arises from the greater command of his own immediate problems, the ability to correlate his discoveries with the facts already known and, perhaps, the contribution of a significant facet to the whole. He realizes, with Aristotle, that "each adds a little to our knowledge of nature, and from all the facts assembled there arises grandeur."

Science spells modern civilization that, ideally at least, ministers to the health, wealth, and happiness of mankind. And it is a justifiable pride of science that what it gains it never loses. From every standpoint the mere pittance man casts to scientific research returns to him many-fold in comfort, in power, and most of all in a broader and more appreciative understanding of his surroundings. Though the walls of ignorance that still encompass him are high, they are being scaled as one generation after another cuts its niches and gives more sure footing to its successor. The horizon recedes as higher altitudes are attained. But, of course, as Huxley expressed it long ago, "in ultimate analysis everything is incomprehensible, and the whole object of science is simply to reduce the fundamental incomprehensibilities to the smallest possible number." Far from depriving nature of its mystery, science affords a sublime picture of its unity and so still more inextricably interweaves human life with the environment.

The contributors to this volume are leading specialists in their respective fields who present for their colleagues and for scholars at large the latest word in recent researches, in refutation of the general notion that accurate, authoritative scientific exposition must be uninteresting. Indeed, the following pages afford an absorbing glimpse of certain significant trends of "science in progress."

LORANDE LOSS WOODRUFF

New Haven, Connecticut,
December, 1944.

Contents

PREFACE *George A. Baitsell* v

FOREWORD *Lorande Loss Woodruff* ix

ILLUSTRATIONS xiii

I. PSYCHOLOGICAL ASPECTS OF MILITARY AVIATION *Walter R. Miles* 1

II. THE PHYSICAL STRUCTURE AND BIOLOGICAL ACTION OF NERVE CELLS *Detlev W. Bronk* 49

III. ENERGY AND VISION *Selig Hecht* 75

IV. CHEMICAL TRANSMISSION OF NERVE IMPULSES *Otto Loewi* 98

V. THE MATHEMATICAL NATURE OF PHYSICAL THEORIES *George D. Birkhoff* 120

VI. THE MAGNETIC APPROACH TO THE ABSOLUTE ZERO OF TEMPERATURE *Peter Debye* 150

VII. THE DRIFT TOWARD EQUILIBRIUM *Henry Eyring* 169

VIII. STREAMS OF ATOMS *Isidore I. Rabi* 195

IX. ADVENTURES IN VACUUM CHEMISTRY *K. C. D. Hickman* 205

X. PRESENT STATUS OF THE VITAMIN B COMPLEX *C. A. Elvehjem* 249

XI. BLOOD AND BLOOD DERIVATIVES *Edwin J. Cohn* 273

REFERENCES 305

INDEX 325

Illustrations

Figure *Page*

1. Army Air Force Thunderbolts.............................. 4
2. Two-Hand Coördination Test used in Classification Centers.. 7
3. Sample results for Two-Hand Coördination Test............ 8
4. Discrimination Reaction Time Test, used by AAF.......... 9
5. Sample results for Discrimination Reaction Time Test....... 10
6. Mashburn Test, used in Classification Centers............... 11
7. Sample results for Complex Coördination Test............ 13
8. Group test arrangements used in Classification Centers...... 14
9. Sample validity results for total preselection test battery..... 17
10. Controls and instrument panels of Flying Fortress.......... 27
11. Drawing of Flying Fortress fighting stations for men and crew 31
12. Round trip in a decompression chamber for high-altitude tests 33
13. Use of oxygen masks in decompression chamber "flights".... 37
14. Preparation for experimental parachute jump by Lt. Col. W. R. Lovelace.................................... 39
15. Drawing of ancient "black care" devils, ancestors of modern "gremlins" 43
16. Flying Fortress with its protective cover of Lightnings...... 45
17. Distribution of calcium in nerve cells..................... 54
18. Nerve fiber of crayfish photographed with crossed Nicol prisms 54
19. Oscillations of potential in region of nerve fiber from which calcium was removed................................ 56
20. Impulses initiated by constant pressure of sensory nerve fiber 58
21. Variation of electric potential on nerve fiber................ 59
22. Impulses passing over nerve fiber from motor nerve cells to muscle 60
23. Nerve impulses from photoreceptor initiated by light of constant intensity 61
24. Effects of decrease of oxygen in nerve cells, measured by oxygen electrode 63
25. Oxygen electrode in position on surface of cerebral cortex.... 64
26. Oxygen concentration in cerebral cortex, measured by oxygen electrode 65
27. Periodic variations of oxygen concentration in cerebral cortex 64
28. First oxygen supply system for high-altitude flight.......... 67
29. Periodic impulses in nerve fiber activating muscle of inspiration 68
30. Oxygen regulator valve used with face mask.............. 69
31. Variations in frequency of nerve action with changes of arterial pressure.. 70

32. Impulses from sensory nerve cell stimulated by gravity...... 71
33. Forces acting on flyer in aerial maneuvers.................. 72
34. Optical system to measure minimum energy necessary for vision 80
35. Diagram of horizontal section through human eye........... 84
36. Chart showing transmission of light through human eye..... 85
37. Diagram showing structure of the retina.................. 86
38. Sensibility distribution in spectrum near threshold of vision.. 87
39. Chart of absorption spectrum of frog's visual purple......... 88
40. Percentage absorption spectra of visual purple concentrations 90
41. Scotopic luminosity at retina compared with visual purple absorption 91
42. Curves showing Poisson probability distributions........... 94
43. Relation between energy content of light flash and observed frequency 96
44. Quantitative result of Parallelogram Law.................. 127
45. Resultant of two equal forces acting at right angles.......... 128
46. Mathematical proof of the Parallelogram Law.............. 129
47. Operations of rotation with square ABCD................. 133
48. Diagram showing group operations of square ABCD........ 133
49. Magnetic moment shown as function of the quotient......... 154
50. Diagram showing procedure of magnetic cooling............ 155
51. Distribution of the 23 electrons in a trivalent Fe-ion...... 157
52. Entropy S as a function of H/T for an atomic weight of Fe-ions 159
53. Entropy values of an atomic weight of Fe-ions as a function of temperature 160
54. Diagram showing elementary cell of $Fe(NH_4)(SO_4)_2 + 12 H_2O$ 161
55. The energy W of the 6 different states of trivalent Fe-ion.... 162
56. Expected decrease in entropy of iron-ammonium-alum under influence solely of H_2O molecules........................ 163
57. Entropy of unmagnetized iron-ammonium-alum as a function of temperature 166
58. Successive states in reaction $AB + CD \rightarrow AC + BD$...... 171
59. Successive configurations in atom reaction $A + BC \rightarrow AB + C$ 172
Curve A. Radiofrequency spectrum of potassium, atomic weight 39 200
Curve B. Radiofrequency spectrum of the H_2 molecule.......... 201
Curve C. Portion of the radiofrequency spectrum of HD, taken in the same way as the H_2 spectrum shown in Curve B.. 202
60. Elementary physics of distillation........................ 210
61. Miscellaneous types of glass pot stills...................... 212
62. Glass falling-film still.................................. 213
63. Early centrifugal still.................................. 216
64. Centrifugal rotor, pick up, and pendulum.................. 219
65. Density comparator 220

66. Rate of travel of distilland.......... 221
67. Photograph of spiral of dyed and undyed oil.......... 222
68. Falling-film still in action, seen by infrared light (Plate I) .... *facing page* 224
69. Infrared absorption and reflection by an oil layer.......... 225
70. Fish-liver oil tank farm.......... 226
71. Commercial centrifugal stills.......... 226
72. Blending and storage tanks for vitamin A concentrates.......... 226
73. Vitamin A, formula, absorption spectra, and properties.......... 227
74. Pure crystalline vitamin A alcohol (Plate II)..... *facing page* 228
75. Plotting ultraviolet absorption curve of vitamin A.......... 228
76. Bioassay of vitamin A.......... 230
77. Cyclic batch still for quantitative molecular distillation.......... 231
78. Elimination curve of vitamin A.......... 233
79. Elimination curve of homologous series of anthraquinone pilot dyes 235
80. Distillation of pilot dyes (Plate II).......... *facing page* 228
80a. Separation of pilot dyes (Plate III).......... *facing page* 236
80b. Samples from commercial distillation of crude palm oil (Plate III) .... *facing page* 236
80c. Distillation curves of vitamin A fish-liver oils.......... 236
81. Master elimination curve of natural vitamin D.......... 237
82. Redistillation curves of natural vitamin D.......... 238
83. General formula for the D vitamins.......... 239
84. "Rat-chicken" structure of high- and low-boiling vitamin D.. 239
85. Distillation map of some constituents of fish-liver oil.......... 240
86. Probable formulae for vitamins A and A_2.......... 242
87. Absorption spectra of whale-liver oil, vitamin A, and kitol... 243
88. Distillation curves of vitamin A recovered from liver of rat.. 246
89. Thiamine hydrochloride 250
90. Riboflavin 251
91. Nicotinic acid and nicotinic acid amide.......... 252
92. Pyridoxine hydrochloride 253
93. Pantothenic acid 253
94. Choline 254
95. Sulfaguanidine 257
96. Succinyl sulfathiazole 259
97. Bacteria (*E. coli*) in cecal contents.......... 261
98. Effect of biotin deficiency on growth of the Hamster.......... 263
99. Dermatitis in chick produced by biotin deficiency.......... 263
100. Biotin 264
101. Effect of nicotinic acid deficiency on chick development.......... 267
102. Para-amino benzoic acid 267
103. a. Blood in capillaries showing red cells.......... 276
b. Blood in Red Cross collecting bottles

104. Relative dimensions of various blood proteins.............. 281
105. Products prepared from Red Cross blood.................. 288
106. Fibrinogen, dry and in solution, and fibrin products......... 292

Science in Progress

Science in Progress

I

PSYCHOLOGICAL ASPECTS OF MILITARY AVIATION

By WALTER R. MILES
Yale University

FLYING, having been for centuries a fanciful, wishful experience of the imagination, has now, in the space of a few decades, become a most practical, effective, and realistic mode of human transportation. The experience of the plane and the planet, the realization of the conquest of time and space by a scientific man-made mechanism is still thrilling in the realm of the imagination. If we had time we might pause to consider some of the fascinating and almost limitless psychological aspects of this new human experience, available as they are for observation to each one of us who has flown through the air. Important to scientific psychology in 1944 as to humanity in general is, however, not the thought but the fact of great distances traveled, great heights reached, national boundaries crossed as if they had no existence in reality, the earth ringed around in a few hours, and man, the object of our study, able to ride at tremendous speed through the air in a powerful ship guided on its mission by his hand. First pushed forward for practical service in the first World War, aviation was developed into general usefulness in the interwar period. In the past five years, first in the prewar emergency and then in war itself, the psychology of aviation has developed particular specialties now of immediate service to the Armed Forces, and, I believe, of ultimate value to the peacetime expansion and development in aviation which will follow victory.

ACCEPTANCE OF HUMAN FLIGHT

Psychologically in the long history of man's development it is of vast importance that a realm can now be directly explored and a type of activity directly engaged in by many people which was formerly accessible only in thought. This suddenly widening sphere for life experience already exercises a strong influence on the development of many human lives and on our total civilization. With the first transcontinental and transoceanic flights of little more than fifteen years ago, flying became possible for mankind in general. From that time onward, a gradual psychological change has taken place in the characteristic attitude of people toward aviation. Although some older folk still think of flying as a dangerous and hazardous undertaking to be avoided if possible, the young now take it for granted, recognizing that, as in the railway or the automobile, hazards are involved but accepting the risk as a matter of course, and not giving it undue weight.

In approaching any study of the psychological aspects of aviation, one must think first of all of this gradual change in the mass human acceptance of flying as part of man's natural right, and at the same time one must not forget the awareness, which continues to persist, of the dangers involved in leaving the solid earth to sail upon an invisible, intangible something which our entire experience says will not support us. Important psychological considerations for personality and behavior are involved in the wish to fly and in the joy of flying. Other factors equally important have their origin in the hesitation or resistance to flying which may be consciously or unconsciously part of the beginning experience. The fact that anyone may fly but that some are better suited to flying than others depends largely on the balance in a given individual between the positive acceptance of flying as natural and desirable and the negative rejection of it as unreliable and unpleasant.

In the mind of many a normal youth the typical adventurous goal in a life on the sea has now been transferred to the air. To fly has become a matter of course to the majority of the young, an accepted mode of activity that is pleasurable, adventurous, favored by strong popular approval, and that may afford a gallant means of serving in the cause of freedom.

"There are three secrets of human flight," says Langewiesche [36]. The first has to do with wings and how they fly; it means knowing what an airplane will do, knowing it with your hands and your brain and your vegetative nervous system. It means keeping your nose down, when logic tells you it should go up. It means identifying yourself with your plane so that its erratic behavior no longer seems unreasonable to you, so that you are content to go where it goes, relaxed, confident, accepting its competence to balance and fly itself on the air as a ship sails or floats on the sea.

The second secret has to do with freedom of motion. It requires your human organism to uproot itself from the solid and stationary earth; to forget the visual cues that predispose you to fly with the ground flat and even below you; to float freely in the substantial air that supports you and your plane in motion so easily and so completely, just as the water does a confident swimmer whether in active motion or carried by the rolling waves.

The third secret has to do with your senses; it means developing a new way of seeing, hearing, feeling, a new sense of touch, and adding all these together so that you "know your way around." When this new sense is fully developed you can fly "deaf, blind, and numb," upside down, by the instruments, by the very feel and sound of wind and air, and by answering to the response all through you, in your joints, your bones, your muscles. It is the unexpected feel of your nervous system as you fly a curve. "I know," says the student pilot, "I get it by the seat of my pants."

Birds have been better prepared by nature than man to learn these three secrets. They get their feelings, as they fly, directly through wings, body, and tail. But man's wings give him no sense of air contact and pressure. He must substitute other perceptions: changes in supportive pressure, and data from vision and hearing. How well he does this determines his success as a pilot (Fig. 1).

The task of psychology in aviation is fourfold: first, to discover the natural talents of men for learning the secrets of flying; second, to develop methods of instruction that can most quickly and effectively impart these secrets to him; to devise ways of protecting him from the special dangers and environmental hazards that lurk in the air, especially at high altitudes; and, finally, to help to sustain in him, in spite of fatigue and the special stresses and frustrations he may experience, the spirit of adventure and courage that is his natural right. These four psychological aids to flying: *selection, training, protection,* and *maintenance,* have developed recently and rapidly. I shall attempt briefly to outline the significance and the contribution of each of the four.

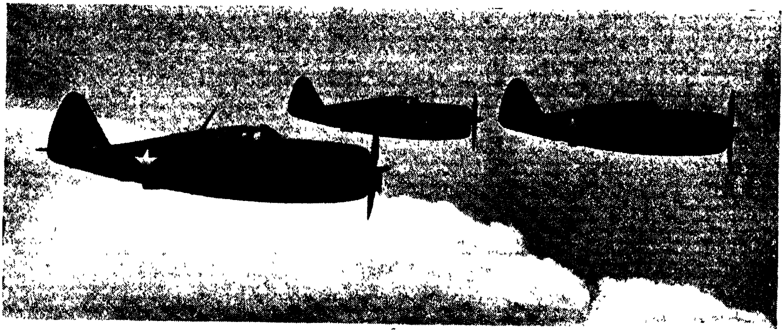


FIG. 1. Pilots with Thunderbolts going through their paces for an Army cameraman. (AAF photo by Hugelmeyer.)

In the years following the remarkable experimental success at Kitty Hawk in 1903, flyer and pilot selection was almost wholly self-selection, depending upon interests and daring. The contract made in February, 1908, for the first Army plane was followed in September of the same year by the Army's first fatal airplane accident. Thus the paradox of progress and its price was demonstrated in military aviation.

The theoretical value of physical fitness and psychological aptitude, in addition to self-choice or military assignment, as a basis for flyer selection, began to be recognized as soon as institutional flying began [11]. The logic of this value was clear: flying was a hazardous occupation, requiring new, complex, and, humanly considered, artificial responses; it had to be learned and practiced under unique conditions of danger and stress. Obviously men in the prime of life and in the pink of condition, who chose the risk, seemed to be the hopeful candidates, but it was clear that their own decision should not wholly determine the matter. There should be objective tests to check their judgment. The wisdom of the decision to exercise medical supervision in the selection and maintenance of flyers was internationally accepted before 1920. Those responsible for military aviation wanted the right men in the planes. This was for the flyer's own sake, for economy in preserving elaborate and expensive equipment, and for fighting effectiveness.

The development of the program of military and naval aviation progressed slowly after 1918 until the late 1930's. Then danger was sensed and a rapid change took place. In 1939 the Civil Aeronautics Administration set up its program for civilian pilot training. As an integral part of this program, a research section was established and a number of research projects were started under the jurisdiction of the National Research Council's Committee on Selection and Training of Aircraft Pilots organized to assist the C.A.A. Thus was initiated the first large-scale effort to develop a rounded all-over objective measure of flight aptitude, correlated with flight training.*

Experimental and clinical psychologists took their responsibility and their opportunity most seriously. Quite a number of them learned to fly. They studied the flying job in detail. They sampled and evaluated the whole range of possible useful types of aptitude, performance, and intelligence tests.

---

* Men active in the civilian work of the Research Council's Committee on Selection and Training have gone on into the armed services. The first chairman of the committee, Dr. J. G. Jenkins, now Commander, in the Bureau of Medicine and Surgery, has a large responsibility in the psychological test program for Naval Aviation. Colonel John C. Flanagan, Chief of the Psychological Section of the Office of the Air Surgeon, was formerly an advisor of the Committee and then an active member. The Committee, now under the chairmanship of Dr. Morris S. Viteles of the University of Pennsylvania, has had as members representatives of Army and Navy aviation, C.A.A., and other civilian specialists.

Objective predictive test results were checked against the success of civilian pilots in training. Measures used in World War I were reëvaluated and sifted on the basis of new data for hundreds of naval aviation cadets and later thousands of army trainees. The motivation in students began to take on the stronger character of an emergency prewar, and then a war, situation. At the same time that the patriotic demands for flyers were increasing, psychological and other scientific methods and techniques for improving selection, training, and maintenance were definitely advancing.

Prewar study advanced job analysis and personnel screening for both ground- and air-crew training. With war expansion selection of men for Army and Navy flying became a first consideration. Physical fitness had always, and rightly, been recognized as primary for fighting flyers. Now standards of psychological fitness were added. Competence was critically evaluated in terms of coördination, intelligence, attention, memory, learning capacity, judgment, emotional stability and integration, traits of temperament, and the total organization of personality. Concurrently with the test selection of men for the air services came their assignment to ground or air crew, and if to the latter, as potentially best adapted for the work of pilot, navigator, or bombardier.

The presifting brought no one-to-one correlation of success in the chosen branch but it is stated by those in authority to have saved expense, man-hours in teaching and in learning, and distress incident to later disappointments. The qualifying medical and psychological examinations had to be limited to the practical minimum requisite for reliability and validity. The batteries now used include what are believed to be the most practical and economical tests of aptitude, intelligence, personality, and temperament. Along with the practical screening the search for improved methods and instruments goes forward.

SELECTION OF FLYING PERSONNEL

Job analysis of aviation always emphasizes four typical aspects of activity requiring excellence in four spheres of human behavior. (1) The necessary aptitude is complex, requiring constitutional superiority developed through interest into primary motor skills of a high order. (2) Intelligence is required at a level definitely above the general average, in alertness and mental organization, logical, realistic, sustained thinking, and practical concrete judgment. (3) Personality traits appear to be most favorable when they result in stability, persistence, good personal relationships, and normal social adjustments. (4) Temperaments are favored when evenly sustained, indicating common sense, with dependability and responsibility as matters of course rather than as remote ideals or stressful monitors.

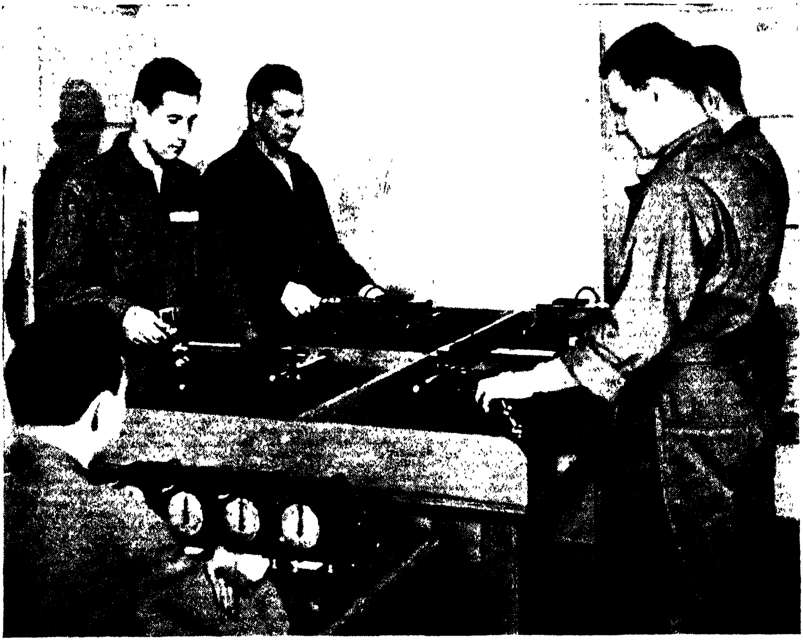


FIG. 2. Four cadets engaged in taking the S.A.M. Two-Hand Coördination Test, as used in the AAF Classification Centers. (Official photograph, U. S. Army Air Forces.)

The specific devices in use and the test items found best for practical examination in the armed services can naturally not now be shown or described in detail. From non-confidential material and from knowledge of the field it is possible to review the types of equipment and the kinds of testing instruments upon which the aviation psychologists have drawn and to illustrate these with examples that have been released for the use of scientists and technologists. The released reports from Army and Navy sources have been augmented by reports from British and Canadian investigations. No secret information can, or should, be revealed.

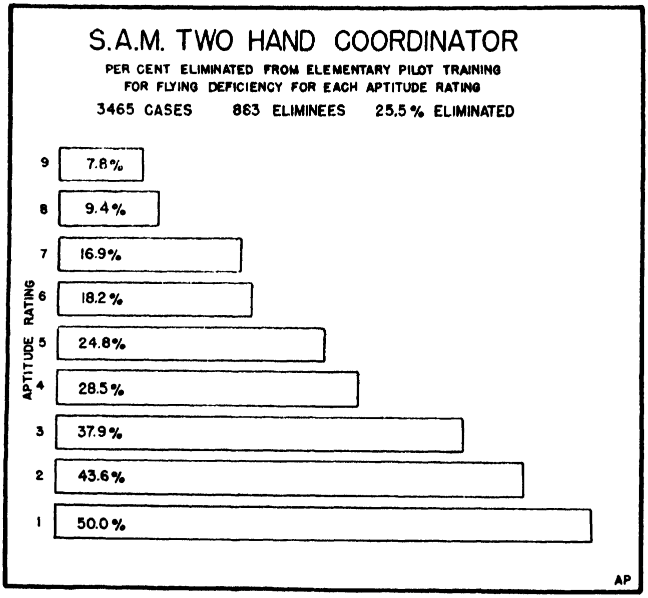


FIG. 3. Sample results for the S.A.M. Two-Hand Coördination Test. (Courtesy of the School of Aviation Medicine, Randolph Field, Texas.)

The various test series so far set up represent planful attempts to measure in sample the gamut of human capabilities in the psychophysiological scales from those most clearly muscular and sensory to the deeply rooted traits of intellect and emotion.

Proceeding from the more to the less objective, the psychomotor reaction time coördination tests are first described. Three of the tests of motor skills currently used in the AAF Classification Centers [44] will illustrate the type of instrument and the kind of psychological reaction which it has.

proved desirable to measure in preliminary military selection.*

The first of these embodying principles of the 1917 Miles Pursuitmeter [45] is a Two-Hand Coördination Test previously used in industry as a measure of freedom from proneness to accident. This apparatus, shown in Figure 2, resembles the head of a lathe which can be moved by the rotation of two handles arranged at right angles to each other, one handle to be turned by each hand. Turning the handles controls the movements of a pointer which the subject tries to keep on a moving target. The target is rotated irregularly by a cam. The score is the amount of time that the pointer is in contact with the target. Each single test is brief, but repeated trials may be given. By the use of identical units, several men can be examined at once. Poor score and accident proneness correlate well on this test, and test scores in repeated trials by the same men show highly similar rank

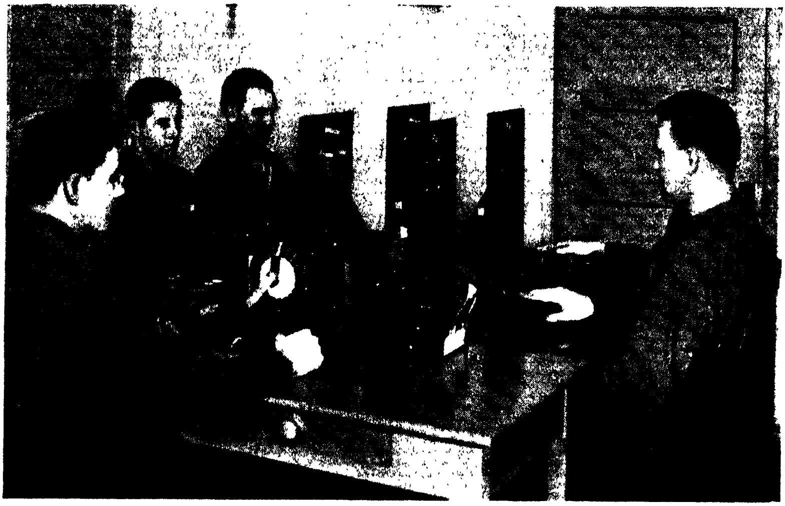


FIG. 4. Aviation candidates taking the S.A.M. Discrimination Reaction Time Test. (Official photograph, U. S. Army Air Forces.)

---

* For these materials the writer is indebted to Lt. Col. Arthur W. Melton, Chief, Department of Psychology, School of Aviation Medicine, Randolph Field, Texas, also see [44].

orders. Test score in two-hand coördination has agreed with ratings (pass-fail) of elementary flight training on several thousand cadet trainees as follows: On a scale of 9 aptitude-rating positions set in terms of standard scores (Fig. 3) only 7.8 per cent of the top or 9 group failed in their elementary flight training. In the lowest or 1 group, 50 per cent were washed out. From group 9 down to 1, there is a fairly regular decline in success.

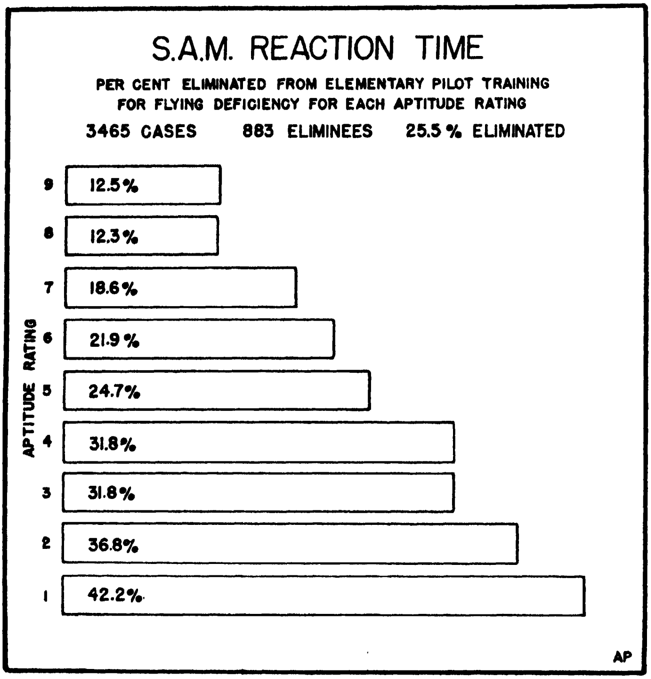


Fig. 5. Sample results comparing score groups and elimination percentages for the S.A.M. Discrimination Reaction Time Test. (Courtesy of the School of Aviation Medicine, Randolph Field, Texas.)

The Discrimination Reaction Time Test is a second unit in the Air Corps psychomotor test battery in current use. This also gives a visual motor reaction time score, but measures a different aptitude from the first described test. Four signal lamps and four response keys are used (Fig. 4).

According to the pattern of red and green lights which appears, the candidate must press one of the four keys. Quick response in pressing the appropriate lever counts toward a good score. Slow response, or worse still, pressing the wrong lever, reduces the score. The purpose of the test is the measurement of speed and correctness of motor response to intermittent visual signals. One examiner can put four men at a time through their paces. Test reliability is relatively good. Validity appears in the practical terms of elementary flight training success (Fig. 5). In top score, group 9, only 12.5 per cent of the men in a sample population

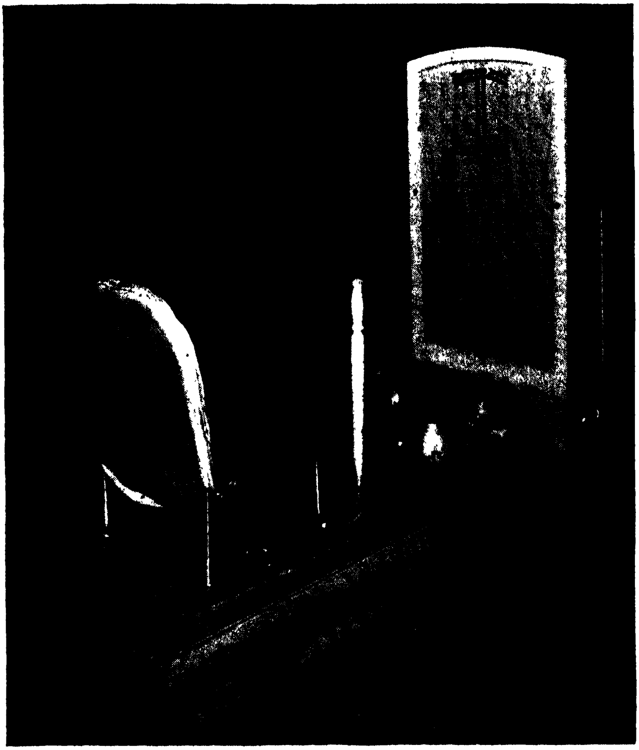


FIG. 6. The S.A.M. Mashburn Test as used in the AAF Classification Centers. Four men each working on a Mashburn unit are tested at the same time in the regular psychomotor examination. (Official photograph, U. S. Army Air Forces.)

failed in elementary training. In group 1, the lowest score group, 42.2 per cent failed. This is less good than the two-hand coördination test as a predictor of training success.

The Mashburn Serial Action Apparatus permits an experience under standard test conditions that resembles in a simplified form the mechanical adjustments of flying. The apparatus (Fig. 6) consists of a cockpit with a stick and a rudder bar and something corresponding roughly to an instrumental panel on which parallel rows of small light bulbs are displayed. This is a modernized and improved version of an apparatus pattern employing airplane controls used and practically modified by Mashburn [39], and employed in the work of the Committee on Selection and Training of Aircraft Pilots. Here, as in the discrimination reaction test, the response is to light signals varying in position. On a vertical panel facing the cadet, are three double rows of lights; a red and a green row form a pair. The curved rows at the top correspond with lateral movements of the stick. The central vertical pair of rows represents elevator stimulus-response lights; the top green lights respond when the stick is pushed forward, and the bottom lights when the stick is pulled back. Straight horizontal rows on the panel show the action of the rudder. Pressing right rudder to the limit illuminates the green light at the extreme right end. The left rudder takes care of the left end. In the neutral position for rudder and stick the central green light will show in each of the three green rows. When the test is to begin, a flashing mechanism simultaneously illuminates three red lights, one in each row. As rapidly as possible the cadet moves the stick and rudder bar so as to match the red lights, each with its corresponding green light. When the matching is complete at the same instant for all three points a new stimulus appears; the flasher moves forward to present three more red lights and so on. A candidate's performance is measured by the number of settings he completes in a specified period of time [54]. Responding in the right direction with proper amplitude, and with both the stick and the rudder bar at the same time, makes for a better score than mere quickness of response first in one axis and then in another, first with the stick and then with the rudder bar. In the practical validity tests of Mashburn Serial Action scores versus elementary flight success the findings are also favorable. Results for two groups of cadets are shown in Figure 7, a Randolph Field group studied in 1931-34, and a Maxwell Field group examined in 1942. In group 1, at the bottom, the failure incidence was 80.6 per cent for the Randolph Field sample and 72.2 per cent for the Maxwell Field cadets. From here on up in score groups 1 to 9 a substantially linear relationship appears.

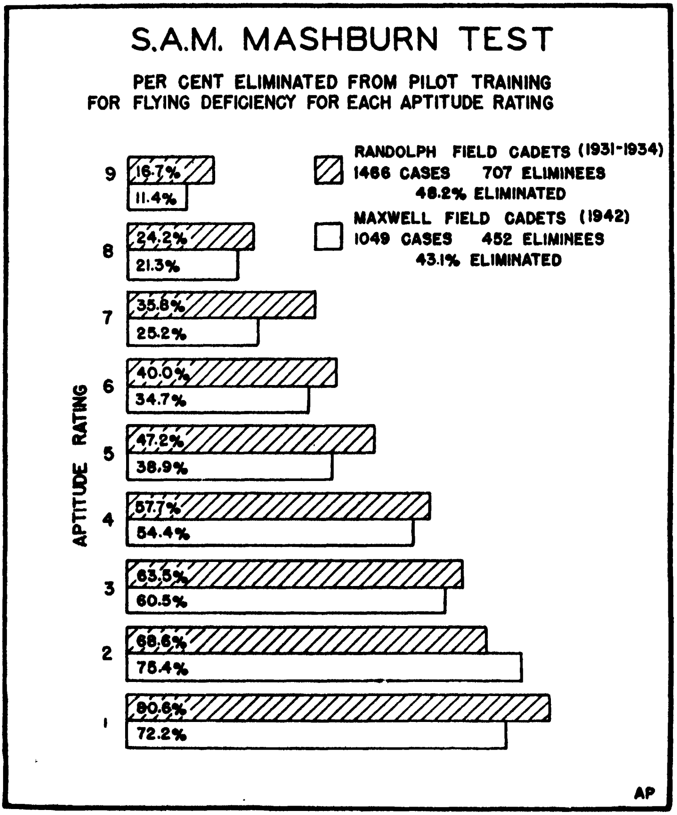


FIG. 7. Sample results comparing score groups and elimination percentages for the S.A.M. Mashburn Serial Action Test. (Courtesy of the School of Aviation Medicine, Randolph Field, Texas.)

The psychomotor classification tests in use in the AAF Centers in July, 1943, included the three already described and three others: rotary pursuit test, steadiness-aiming test, and finger dexterity test. They are called the S.A.M. tests, since they were selected, or developed and arranged for administration and validated at the School of Aviation Medicine, Randolph Field [54]. A rating made up of the combined scores for these six S.A.M. tests gives a definitely favorable predictive measure. However, this is not used by itself but in combination with scores from all the objective psychological tests, including paper and pencil tests for intelligence, special aptitudes, biographical inventory, and others.

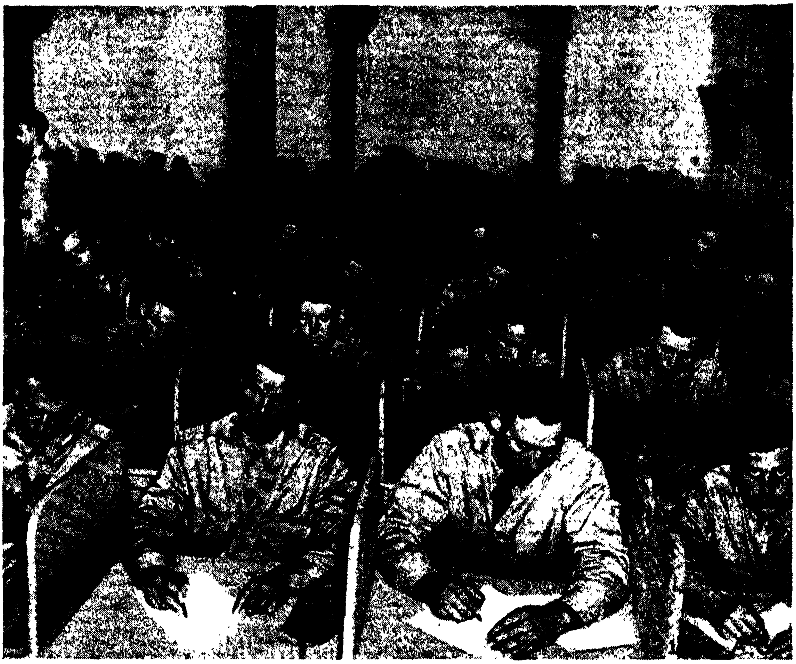


Fig. 8. Arrangements for giving group tests in the AAF Classification Centers. (Official photograph, U. S. Army Air Forces.)

The intelligence, mental alertness, and aptitude tests employed for the selection of flyers are up-to-the-minute versions, expressed in aviation terms, of the best verbal, mathematical, and reasoning tests and test items. The battery is given by the group method under arrangements such as those illustrated in Figure 8. Non-verbal tests, especially measures of spatial relationships, fundamental mechanical concepts, quickness and accuracy in observation and in following directions, visualizing and drawing conclusions from visualization, flexibility of attention, making logical deductions and formulating good judgments from realistic, often complex data, all of these constitute the mental elements and factors of intellect, basically important for flyers as for all individuals who must act promptly, responsibly and adaptably to constantly changing, always dangerous, and generally complex situations. The types of items are well known. They include effective samples from the standard test categories suggested by Spearman's G and specialized factors, Thorndike's CAVD, Thurstone's primary mental abilities, Terman's Binet scale, and all the vast selective paper progeny of Army Alpha and Beta. A few examples will suffice. The following were reported by Flanagan [30]. The correct answer for each item is here printed in capitals but quite obviously would not be so designated in the actual test sheet supplied to candidates.

(3) As it rounds a sharp curve an automobile goes off the road, turns over completely once (landing on its wheels), and skids onto a smooth ice-covered lake. Under these conditions the car is most likely to:

3–A Roll forward without turning over again.
3–B Turn over and over.
3–C SKID IN A RELATIVELY STRAIGHT LINE.
3–D Skid in a wide circle.
3–E Roll backward without turning over.

Answered correctly by 78 per cent of the passing group and 50 per cent of the failing group in elementary flight training.

(4) Bombing attacks are usually made at night chiefly because:

4–A PURSUIT PLANES HAVE GREAT DIFFICULTY LOCATING THE BOMBERS.
4–B Radio detection devices do not work well at night.
4–C Fires started by the bombs can be used as targets.

4–D They prevent the enemy population from sleeping.
4–E They are as effective as daylight attacks but are less dangerous.

Answered correctly by 66 per cent of the passing group and 37 per cent of the failing group.

(6) The formula $F = \frac{9}{5} C + 32^\circ$ expresses the relation between degrees Fahrenheit and degrees Centigrade. 20° Centigrade equals what Fahrenheit temperature?

6–A 96.6° 6–C 212° 6–E 68°
6–B 4° 6–D 36°

Answered correctly by 90 per cent of the passing group and 36 per cent of the failing group.

Completion of the battery of paper-and-pencil tests requires one day of the cadet's time and the S.A.M. psychomotor tests an additional two hours. Each test considered by itself has proved in some degree to be predictive of flying success. A composite score based on the entire battery of psychological tests compared against per cent elimination from elementary pilot training, for a fairly large sample of 29,069 cases, gives the results embodied in Figure 9. Here, aptitude rating 9 group shows only 6.4 per cent failure while 1 group shows 81.5 per cent of washouts; from ratings 9 to 1 the elimination percentages increase quite regularly. Melton [44] states that the AAF "now rejects for training any candidate who has an aptitude rating of 1, 2, or 3." Let us note what this will mean with reference to the sample of cadets represented in Figure 9. Training was undertaken on a total of 29,069 men; 18,754 got their wings, 10,315 failed, and the attrition was 35.5 per cent. If we reject the first three aptitude rating groups, which include about 17 per cent of the sample population, 4,888 men would be eliminated, of whom 3,152 or 64.5 per cent proved failures after training, and we should also eliminate 1,736 passers, which is 9.2 per cent of the successful men. Aptitude rating groups 4 to 9 inclusive contain 24,181 candidates, of whom 7,160 or 29.6 per cent failed, which is 83.4 per cent of the attrition rate found for the entire sample. This is a considerable gain and on a purely monetary basis worth much more than the cost of the screening process. In terms of human values the reduction of washout rate with its consequences in morale is of great importance. But psychological preselection is not a simple matter. Many considerations enter when determining the position of the cut-off score or aptitude rating below which candidates will be rejected. In the sample presented a point has been chosen which will probably eliminate about twice as many potential failures as potential successes. Even though flying personnel is urgently needed, the AAF administration chooses not to take the training risk when the chances of failure are 2 to 1. Further improvement in the preselection battery of psychological tests for airmen may be made but the score range cannot of itself be expected to specify the rejection point.

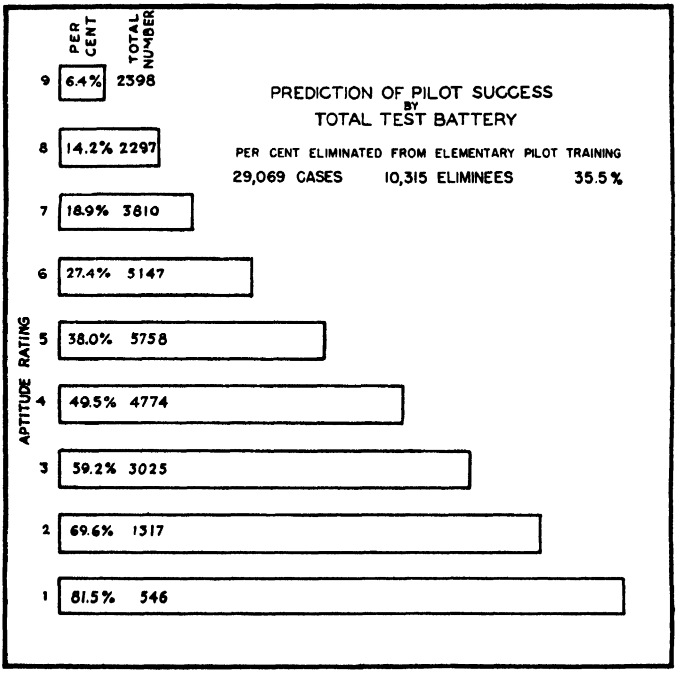


Fig. 9. Sample validity results for a total psychological preselection test battery used in the AAF Classification Centers. (Courtesy of the School of Aviation Medicine, Randolph Field, Texas.)

Psychologists have tended to be critical of personality and temperament tests. Perhaps too conscious of the lesser adequacy and predictive value of such instruments, when compared with aptitude score usefulness and dependability, they long hesitated to put them forward for what returns they might bring into the general rating pool. In the conclusions of psychologists in 1919, as to test function in pilot selection, personality tests got scant mention. Civilian pilot testing came to rather negative conclusions regarding the usefulness of test material of this type. Clinical interpretations were needed to make the data useful. Personality test inventories as unit measures have not contributed to total selection ratings in any substantial way. But item analyses of these tests (biographical inventories) have brought to light predictive material and the test patterns can give good hints to those who know how to interpret them with respect to the needs of the various types of men who learn to fly.

Finally in 1942 a systematic program was set up in Psychological Research Unit No. 1 of the Army Air Forces [53] to develop tests of emotion, personality, and temperament as part of the preselection of Army air crews. From intimate and detailed observations and discussions on the ground and in the air, broad areas were defined for personality research. These included categories for investigation such as resistance to distraction, freedom from fear of physical danger and of failure, resistance to criticism, endurance of fatigue and discomfort, consistency of performance, resistance to confusion in performing complex tasks, motivation and interest, social adjustment, emotional stability, and finally a total mode of perceiving, interpreting, or judging situations, which is based on a constellation of emotional habits and forms a typical "apperceptive pattern." Construction of tests is under way in Unit No. 1 to assay personality within these categories and to determine the relation of traits or trait groups, thus measured, to success in air-crew or ground-school training and later combat duty. A keen analysis of the problems and assumptions involved in their study has been made by the personality psychologists and their chief, Lt. Col. Shaffer. These investigations have already resulted in the development of techniques for studying a number of trait groups or habit constellations. One of these is an interest test designed to discriminate patterns of likes and dislikes for sports and hobbies to be validated for the three flying specialties (pilot, bombardier, and navigator) and ground crew. Another is the control confusion test designed to measure resistance to confusion and breakdown in the performance of tasks that change rapidly or that become increasingly difficult or both. The apparatus for this test

> consists of seven airplane controls each of which has a number of contact points permitting the correct setting to be changed by the examiners. The task of the candidate is to stop the hand of the clock in front of him, getting all seven controls correctly set at the same time. The correct setting for each control is changed frequently during the test so that the subject has constantly to adjust to a changing stimulus situation. A record is kept of the length of time during the test that the subject has all seven controls set correctly, as well as a separate record of the time each of the controls is set correctly. In order to gather additional results concerning the performance of this task, observational data are obtained. Emphasis is placed on the subject's general manner of approaching the task, on the amount of confidence he displays in the situation, on the precise manner in which he operates the controls, and on any additional features which appear significant to the experimenter.

Including these examples, 63 personality and temperament tests were developed by Unit No. 1, prior to June, 1943. In all of these studies, the clinical point of view was maintained that a man performs his task always as a whole individual. Therefore tests of emotion, personality, and temperament cannot be separated completely from each other or from intellectual, perceptual, and coördination tests.

Test batteries in practice have been described by Kellum [35] and others. One such battery of three tests, validated by the performance of more than 3,000 cadets and used for the selection of candidates for naval aviation training, was made up of paper-and-pencil tests which could be administered in large groups and scored electrically. This type of battery has found a place in the pressure program under which selection is now carried on in the United States Navy. First in this particular series comes an elimination test of intelligence. Only those who succeed above a critical point may go on to the second test, a mechanical comprehension test, and to the third, composed of a biographical inventory of highly valid items. From tests 2 and 3 the candidate is rated in one of five aptitude groupings, the attrition rate from which can be fairly accurately predicted.

In England, in Canada, and in the United States it is usual for military or naval selection to be carried on as a joint responsibility of psychologists and psychiatrists. The emphasis of psychiatry in the early stages of progress was chiefly on the negative side, i.e., the sifting out of the men who were judged unlikely to stand up under the stress of combat flying. The psychologists' goal has been the prediction of the likelihood of success. The combined results, which bring together the data from the most reliable test series with insights from intuitive appraisals by experienced clinicians, have given and still continue to give the best indicators for prediction.

General Arnold once remarked to the writer, "Big men for big planes and little men for little planes," which, apropos pilot selection, is a better line than the one about round and square pegs. Preselection is aimed at keeping men from getting up against or becoming responsible for too much airplane. This is a matter peculiarly critical in military aviation. And it is less simple than it sounds to see who are the big and who the little men. During the last three years military psychologists have had a unique opportunity to explore the problem. Never before have so many been in a position to gather so much in the way of controlled scientific data on potential flyers. A considerable measure of success has been achieved; a good many tests have been proved to have respectable validity as measured against the pass-fail criterion of elementary pilot training. The attrition rate has been sharply reduced and the probabilities for training success are now known for various test battery score ranges. But these scientific expectancies apply to groups of men, not to the individual. The man who at all costs is determined to fly may fly, if not a "big plane," at least a "little plane," in spite of his poor showing on physical examination "64" and all the psychological aptitude battery, provided he does not come up against the "contrary-minded" instructor. At the other end of this scale is the individual who, in the examination and tests, falls well within the top group but in training proves not to have what even the "little plane" takes.

TRAINING FOR MILITARY AVIATION

After selection, training offers the key approach to skilled activity. "Flying is beset with danger, and skill has been bought at the expense of safety," said Lt. Col. Leonard G Rowntree, then Executive Officer of the Medical Research Laboratories of the A.E.F., in reviewing accident casualties in 1917-18 among military flyers. Then, as now, every training accident was begrudged as wasteful and somehow to be made avoidable. The figures for current casualties are of course not available. But those of the first great war illustrate the problem and indicate a trend that psychologists, technologists, and all aviation specialists are seeking to diminish. In the A.E.F. [8] one quarter of the 2,034 flyers and aerial observers, distributed in the zone of advance and allied squadrons, were reported killed in accidents prior to the Armistice. Almost half this number were victims of accidents in training overseas. It was a smaller number than this (169 to be exact) who were reported killed in combat.

Our newspapers in 1944 not infrequently carry items recording airplane accidents of which the following is a sample:

> Deputy Sheriff ———— reported today that two four-engined bombers collided in flight about 17 miles northeast of here killing all aboard both planes. The Sheriff said he talked with two witnessses who saw the two planes collide and that they reported one of them exploded before plummeting to earth. Both planes burned. Four-engined bombers usually carry a crew of 10 men.

What was the cause? Who knows or who can unravel the mystery, with all on board killed and both planes burned, unless there may have been a radio report shortly before the collision? Probably this very costly accident will be charged off under the well-known but poorly defined caption, "pilot error."

We cannot eradicate accidents but we may, by appropriate methods, considerably reduce their number. Relative improvement has occurred through the combined efforts of science and practical experience. Flying accidents are receiving detailed, intensive study from medical specialists, psychologists, and engineers and so our data from expensive and fatal experiments are made effective for those who must carry on. Systematic improvement of training is the main avenue through which flying safety can be approached, while at the same time, skill, power, and efficiency are built up for combat aviation. Selection has raised the fighting aviator's stamina level and done much to eliminate nervous instability from service groups. But no psychologist, flight surgeon, or commander will be satisfied until avoidable accidents, including those due to all personal factors in trainer and trainee, are reduced far below the present incidence rate. Training in aviation is not a culture course. It is a technique for service and survival. Its goal is highly skilled learning, which involves deeply ingrained habits of automatic response to take care of routine situations and a wide range of less used but essential possibilities to be available, with no fumbling of judgment, as the working capital of the aviator in moments of surprise and emergency. Instructors and other flyers do well to check-fly each other objectively and helpfully. In the old days more than one pilot said, "My instructor was killed doing what, years ago, he told me never to do."

Psychological factors in flight training are now actively explored. The personal and the technical elements on both sides are subject to survey leading to improvement in method and content. In earlier years the military aviation training divisions have said, in effect, "Give us the right men to train and we will look after training them. We don't want to train poor men but we know how to train good men." In a sense this is true and all things considered, training, even before any very systematic analysis of it had been attempted, was surprisingly good. Instructors have been enthusiastic, cadets highly motivated, and engineers have done their part by producing reliable training aircraft; yet the elimination rate for elementary flight training has been high [30], in the AAF from 1924 to 1941, ranging from 60 to 35 per cent, and numerous accidents have occurred and still occur. With acceleration in the training program the value of even slight improvements in detail or over-all procedure has increased as a matter of simple arithmetic.

Practical difficulties have had to be overcome. Training seems most naturally to take place chiefly in the plane, and the typical training plane is arranged to carry one student and one instructor. The opportunity for the study and improvement of flight instruction has therefore been limited. As a rule, each instructor has four or five students at a time, seldom more. The close man-to-man relationship that must exist, compatibility between student and instructor, is highly important. As far as possible, students are assigned to instructors in whom they have confidence and, if possible, whom they admire. Personal antagonism or non-adaptability between student and instructor can to a considerable extent be avoided if basic personality traits are taken into account. It is almost needless to emphasize the wide range of individual differences in instructors as well as in students. Instructor effectiveness will tend to vary. Instructor training in uniformity of procedure based on sound principles cannot improve general practice unless procedure is standardized. Different instructors will emphasize different methods and maneuvers and may unintentionally slight or ignore important elements. Inspectors also differ, and it is to be anticipated that instructors will reflect the influence of inspection awareness. All of this is natural enough in human terms, but in a nation-wide pressure-program the defects involved in individualized instruction must be as far as possible eliminated, the excellences preserved and generalized.

In the C.A.A. research program, under the direction of Dr. D. R. Brimhall, a beginning was made in the study and analysis of flight instruction. Manuals were constructed and "patter books" were compiled [20, 21] describing, in practical detail, basic flight maneuvers. For the first time in aviation an effort was made, through research undertaken by Dr. E. Lowell Kelly of Purdue University, to find out how flight instructors actually do their job of instructing. Radio equipment for transmission from plane to ground was installed and recordings were secured of conversations between the instructor and the student during instruction flights. Many hours of actual teaching by different pilot instructors in different types of aircraft and representing different parts of the flying course were recorded, duplicated, and submitted for detailed study, discussion, and revision. Great differences between instructors were found. There were wide variations in the types of explanations accompanying a given maneuver, also in the correctness of the explanation itself from the standpoint of aerodynamics, and there were large differences in the utility and adequacy of the vocabularies used. In combined samples of ten hours' instruction by each of four instructors a total of 500 different technical terms or phrases appeared, many of them poor equivalents for the clearest, best expressions. Each man had his own unique vocabulary, a sort of personal patter, self-developed. Few descriptions of maneuvers which could be considered complete and correct were found in the records. It also appeared that a large amount of the instruction given in the air could be given as well or more profitably on the ground. Too frequently complicated aerodynamic explanations of a maneuver were given to the student while flying with the result that he was incapable of giving his best attention and comprehension to either the flying or the explanation. The instruction from the plane recorded directly on the ground made an excellent beginning from which, through a series of practical revisions with repeated rehearsals in flight, the C.A.A. patter books finally emerged.

The flight-tested instruction techniques of the patter books and manuals are clear and straightforward. They embody the best of the methods of the most successful instructors in the words and phrases that have proved most effective. A sample paragraph from a section on taxiing in *Elementary Flight Maneuvers* (C.A.A. Bulletin No. 21, February, 1943) will illustrate:

> In taxiing we want the tail of the airplane to remain on the ground at all times, so while we are heading into the wind, the stick should be held slightly back of neutral position. Thus the wind will force the tail down. When taxiing with the wind, since the wind will usually be traveling faster than the airplane, the stick should be held slightly forward of neutral position. Thus a sudden gust of air striking the elevators will force the tail down and prevent the wind from getting under it and upsetting the plane. Your hand should at all times rest lightly on the throttle. In taxiing the throttle should be used smoothly and only enough to keep the airplane moving slowly, about as fast as you walk. Hold the stick between the thumb and fingers of your hand. During the taxiing process, follow an "S"-shaped course occasionally so that you can see what is directly ahead of the airplane.

Thus the instructions in simple direct sentences tell what to do and why to do it. This is at once a method of procedure and a provision for safety. All of the usual routine maneuvers can be stamped in as reactive habits.

Methods for solving novel problems and executing the less usual maneuvers have also to be taught and learned. During the training period nothing should be left to trial and error that can be met by analysis and synthesis. The flyer can learn to analyze sample situations and to combine conscious thought with common sense and good intuition. Thus the instructor prepares him for an approaching stall:

> Vision [he will say] which has played so large a part in your previous maneuvers is of little help in sensing an approaching stall, except in noting the airspeed indicator. Rather you have to learn to "feel" a stall; though you won't be able to see a stall approaching you may be able to hear it. Learn to sense a stall by the sound and the feel of your ship. As the air speed decreases, the first symptom of a stall is that the controls, both the stick and the rudder, begin to feel loose or sloppy. That is, they move much more easily at cruising speed. Also you will discover that normal movement of the controls results in much less response on the part of the plane than in normal flight. As the stalling condition gets nearer and nearer, the controls become less and less effective. First you will notice that applying aileron pressure no longer enables you to bank or to pick up a low wing.

Instruction of this kind not only aids learning at a specific point, and speed of mastery of this particular problem; it also gives a standard method of approach to all problems, creates a professional attitude, and tends to build morale. Gradually the trainee learns the value of book study, of note keeping, and review, as preliminaries to exercises in the air. Like the well-coached football or baseball player or the highly trained industrial technician, he is finally able to previsage every possible experience and so his effectiveness is increased in meeting the real situations as they arise. The psychological value of conscious progress, of reaching one by one successively higher levels of competence and of suitable recognition of good work, is known to the instructor and acted upon in the routine curriculum.

Primary and advanced training are both psychologically similar to elementary instruction except with respect to complexity. The requirements at the advanced level (Fig. 10), make tremendous demands on the organization of routine procedures, on depth of insight, and on breadth and functional flexibility of judgment. At these later stages of training the plane presents the technical problems of a more powerful engine, faster flight, more complex controls to regulate, and a more elaborate instrument panel. The meanings of many more instruments and instrument readings have to be learned.

Mixture control, spark control, primer, wobble pump, tail wheel lock, landing wheel switches, and other elements enter the picture. The pilot must now acquire the habit of reaching for any one of the controls automatically. In combat flying he will not be able to spare vision for the direction of all his hand movements. The secondary plane will be heavier than the elementary training plane and will fly differently. The pilot must, as quickly as possible, learn to make the plane do precisely what he wants it to do. In achieving this level of mastery he must learn to regard himself as the pivotal center about which the plane revolves and the plane as an extension of himself. Anticipatory study on the ground with a new plane in the early stages of acquaintance and especially

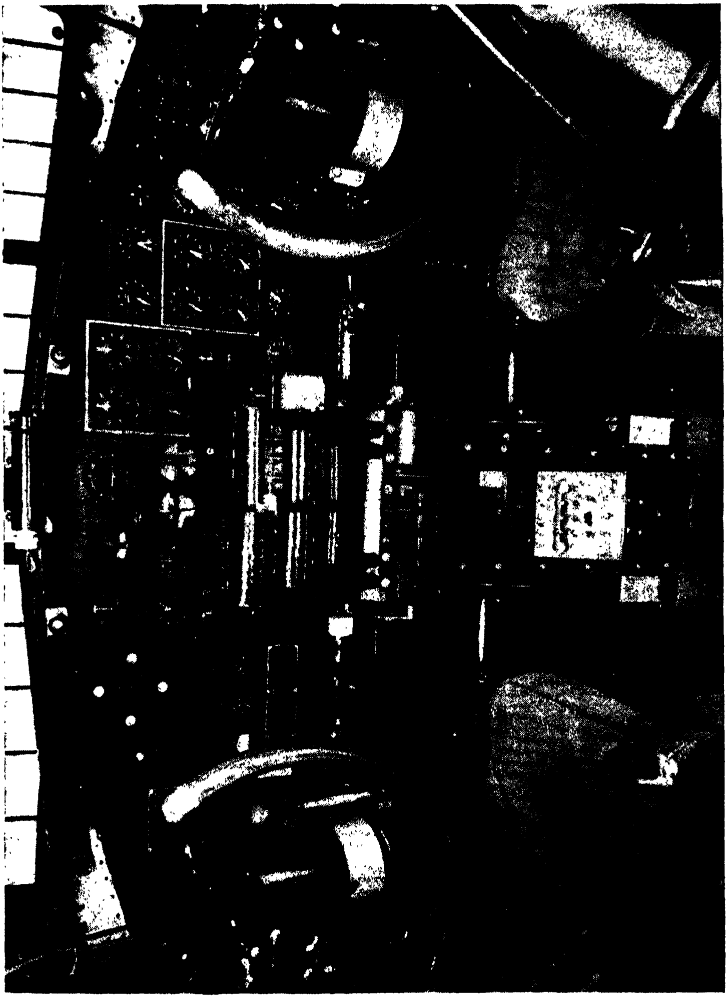


FIG. 10. Some of the controls and instrument panels of the Flying Fortress. The pilot's seat is at the left, the co-pilot's place is at the right. Through training the pilot must set up an elaborate system of response habits which, on the psychological side, is the complement for the instrument complexity indicated in this illustration. (Courtesy Boeing Aircraft Co., Seattle, Wash.)

when aided by modern training films, nets more than an equal period in the air. Manipulation, visualization, and finally flying experience when accompanied by well-formulated indoctrination give the best results.

OBJECTIVE EVALUATION OF TRAINEE PROGRAM

In the past the instructor or inspector was the judge of the student's success in performing the maneuvers involved in a particular flight lesson or check-up. Different instructors and flight inspectors might have diverse standards and methods for rating student performance. Quite dissimilar skills are at times emphasized as the significant criteria of flying excellence. In civilian training, inspectors' ratings have yielded low intercorrelations. It was hardly to be expected otherwise. Something in the way of objective measurement is essential before people can achieve or approach uniform standards.

Objective records of what the student does with the plane can provide a basis for more accurate and reliable grading of learning and competence than subjective observations; when used for comparison with instructor and inspector judgments they can lead to improvement and refinement in estimates. Graphic records can assist in the analysis and correction of faults. Recently it has been announced in the public press that the C.A.A. research program has developed, under the leadership of Dr. Morris S. Viteles, an accurate objective method of plotting plane performance by means of a standard flight recorder. By use of an accessory instrument panel of special design a movie camera is arranged to photograph the way the student handles his controls, at the same time that a record is made of the actions of the plane. In conjunction with the photographic records a magnetic wire recorder, also carried in the plane, enables the instructor or inspector to dictate a full account of his impressions of the student's performance. Analytical comparison of the comments and the photographic records will bring correction and improvement of instructor and inspector judgments, and provide common standards.

In the brief survey here outlined an attempt has been made to indicate the way in which the five essential steps in job instruction can be taken in aviation training. The steps are as follows: 1. Preliminary formulation of what is to be taught, the lesson aim. 2. The selection of material and preparation of method of giving the lesson. 3. The demonstration and instruction presentation to the trainee. 4. Tryout and practice by trainee. 5. Evaluation and review. How well has the lesson been learned and how well is it understood. It has not been uncommon for aviation instructors to eliminate or slight steps 1 and 2, concentrating on the others which are usually given in the plane in the air. However, good instructors have always used all five steps and checked on them. In total war they become the supportive structure upon which effective teaching and training rest. The real pay-off should come in reduced accident incidence and increased combat efficiency.

The individual training of the airman is not enough. This must be supplemented by crew training so that the pilot, co-pilot, bombardier, navigator, gunner, radiomen, and flight engineers all perform as an integrated unit. It is always efficient team work that counts. In Figure 11 a Boeing artist has depicted a Flying Fortress in combat in the South Pacific. Parts of the "skin" of the great bomber have been cut away to show the crew in action. In the nose, the bombardier has just released a bomb, the navigator and engineer are busy with their "fifties," the pilot is shown watching for enemy planes. The co-pilot sits at his right. The top turret guns, manned by another engineer are ready for action; next the bomb racks are shown, and then another top gun operated by the first radioman against targets coming in from above. The second radio operator mans the ball turret against underneath attacks. The two waist gunners are ready to use fast-firing .50 caliber machine guns. The tail gunner is shown kneeling, waiting for an enemy plane to come within range of his twin "fifties." An intercommunication system joins the eyes and trigger fingers of all the crew members so that they coördinate. The Fortress can "take it," and so can the properly trained Fortress crew. The crew plus the flying gun-platform and equipment make the Flying Fortress effective for military aviation.

PROTECTION FOR HIGH FLYING*

The atmosphere in which we live and breathe is uniform in its composition around the earth since it is kept well mixed by the upward and downward currents which result from the unequal heating of the earth's surface. In terms of dry air the constituents are oxygen 21 per cent, nitrogen 78 per cent, and other gases 1 per cent. The ability of oxygen to combine with and saturate the hæmoglobin of the blood depends on the pressure of this gas in the air and the air pressure depends on altitude. At sea level where the air pressure is 760 mm. Hg. the oxygen partial pressure will be 1/5 as great, or 152 mm. Hg., since the ogygen makes up 1/5 of the air. At the sea level this partial pressure of oxygen provides for normal saturation in the blood at 95 per cent or above. As the flyer leaves the ground and climbs to higher and higher altitudes the air pressure falls progressively and accordingly the oxygen saturation in the blood decreases. The effects of the oxygen decrease on the nervous system, measured in terms of the flyer's operating efficiency, are detectable but not very pronounced at altitudes between 7,000 and 10,000 feet. At 18,000 feet the air pressure, reduced now to 1/2 that at sea level, can scarcely maintain better than 70-75 per cent saturation in the blood, an amount which for most men is definitely insufficient for normal psychological and physiological functioning. Collapse or even death from oxygen want has resulted in some men flying to altitudes of 24,000 to 26,000 feet even when the plane was at this altitude for as short a time as 10 to 15 minutes. This was the shocking, unanticipated fate of some of the first balloon flyers.

---

* In October, 1942, the National Research Council, in its Division of Medical Sciences, set up a Committee on Aviation Medicine under the Chairmanship of Dr. Eugene F. DuBois. This group, with its five sub-committees, has been continuously active on problems concerned with the protection of flying personnel and has coöperated with the Army and Navy and their various research laboratories having to do with aviation. Results of this war research are still, October, 1944, restricted from publication.

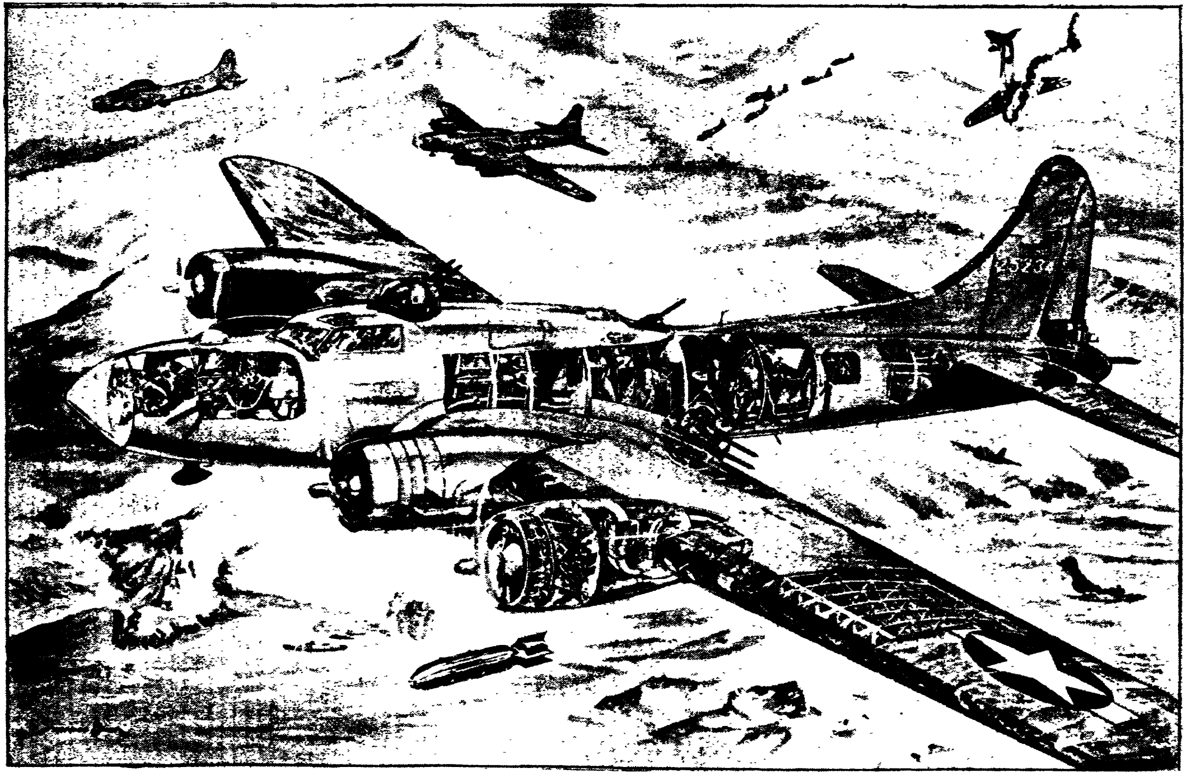


Fig. 11. A drawing indicating the fighting stations of men who make up the crew of the Flying Fortress. (Courtesy of Boeing Aircraft Co., Seattle, Wash.)

If in place of breathing free air a man breathes pure oxygen from an "oxygen bottle," cylinder, or other source through a close-fitting mask which covers the nose and mouth, the saturation in the blood, while at sea level, will still be normal. There will be little or no change in comparison with breathing the 21 per cent oxygen in the atmospheric air at sea level. But with the 100 per cent oxygen taken by mask the flyer can go to an altitude of 39,300 feet where the total air pressure is only 152 mm. Hg., 1/5 atmosphere, and still have normal oxygen saturation, because now all this pressure is oxygen pressure. The critical need for supplementary oxygen was clearly realized in World War I. At that time the need was usually met by having the flyer breathe oxygen from a rubber tube attached to a pipestem held in the mouth. This has proved to be a treacherous method of giving oxygen at high altitudes. Atmospheric air is taken in through the nose and this greatly reduces the partial pressure of oxygen in the lungs. As late as 1935 in a balloon flight to a height of 32,800 feet the crew members died from oxygen lack. They were taking oxygen by mouth tubes but had not realized the necessity of restricting nasal breathing by nose clips or other means.

The problem of supplementary oxygen supply, including comfortable efficient breathing appliances, for flying personnel engaged in military or civil aviation, as now practiced, is protection problem No. 1. Primarily a physiological problem, it presents certain psychological aspects and is best studied by a dual approach. The standard locus for these studies is a laboratory equipped with a decompression chamber (Fig. 12), in which a wide range of altitudes can be simulated, and groups of subjects observed and measured under easily controlled conditions. In the decompression chamber indoctrination of personnel for high-altitude flying can be efficiently carried forward parallel to the investigation of altitude problems. Lessons and demonstrations of physiological and psychological functions as they correlate with oxygen needs can here be given with appropriate practical emphasis. The effects of milder degrees of anoxia can be

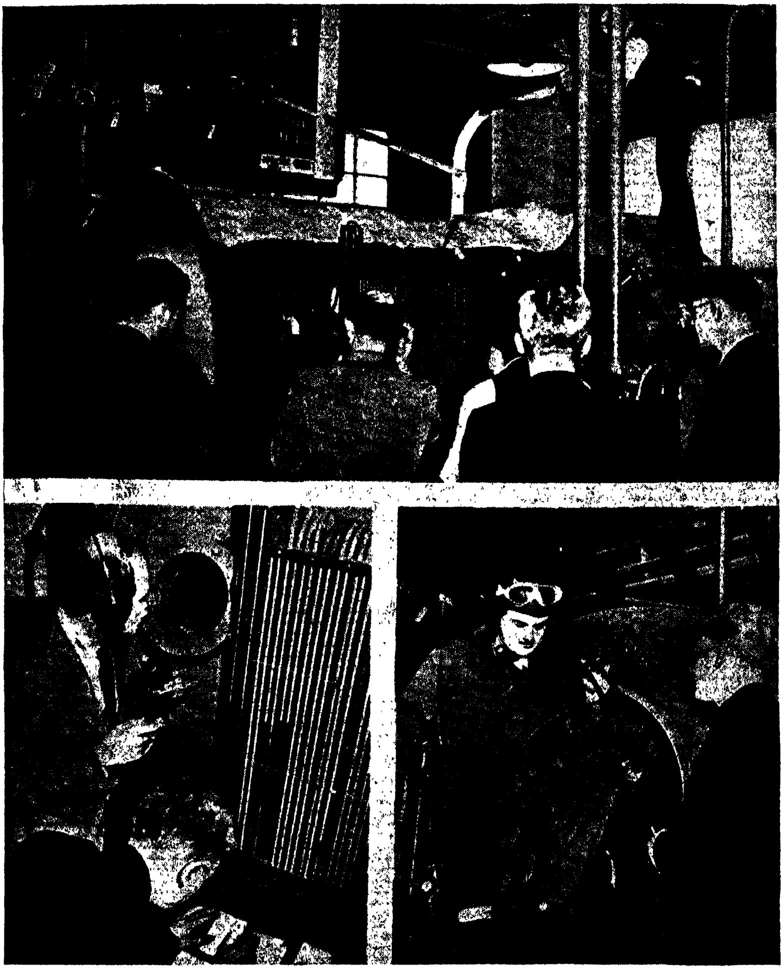


Fig. 12. Upper, a decompression chamber suitable for making high-altitude protection tests on flying personnel. Lower left, observers follow the progress of the training through thick windows in the chamber, communicating by telephone with those inside. Lower right, back on the ground again. A round trip to 35,000 feet in the chamber takes about one hour, much less time than is required in a real airplane. (Courtesy of Boeing Aircraft Co., Seattle, Wash.)

experienced, the favorable, compensatory effect of the oxygen equipment can be shown and its intelligent use routinized.

Anoxia, the state of insufficient oxygen supply to the organism, presents a subtle, psychophysiological danger. The effect comes with no anticipatory alarm signal to a person sitting quietly. The symptoms appear gradually and insidiously in body and mind. Sensory and mental capacity slowly but definitely decline, vision and hearing become less acute, muscular responses, characteristically delicate, are now clumsy, attention weakens, and judgment wanes. The changes in behavior are observable by others not suffering similar oxygen decrement, yet they may occur unnoticed by the person experiencing them. Under whatever conditions the loss of oxygen takes place, i.e., whether in actual flying, in a decompression chamber, or through the introduction of a changed percentage ratio of nitrogen to oxygen, the psychological results are the same if the rates of ascent and the times of exposure are equal. McFarland's (41) results show an increasing scale of sensory, motor, and mental deterioration with gradual loss of oxygen. Fine motor reaction as in handwriting, visual motor coördination in the reaction of color naming and in simple routine decoding, and various acts of attention and memory all show decline in these investigations. Handwriting was impaired at 14,000 feet and above, with marked deterioration in this common expression of muscular habit and control, following rapid ascent to 16,000 and 18,000 feet. The recognition and naming of 100 color squares in chance order showed especially significant change in rate and correctness when oxygen deprivation was followed by a rapid ascent to 12,000 feet and at 14,000 feet and above following a slow ascent. Decoding gradually required more time for equal amounts of work, and errors were increased until a statistically significant accomplishment loss was registered at 12,000 feet and above. The memory tests showed clear decrement after 14,000 feet but with wide individual differences, probably correlated with the type of material to be recalled and power of compensatory resistance to loss in the individual. Recent studies on anoxia by means of visual threshold tests have indicated significant changes even below an altitude of 10,000 feet.

Many researches are going forward, whose objective is a more precise understanding and definition of immediate and prolonged oxygen starvation. It is now definitely known that high-altitude flying can be engaged in to such an extent as to result in a chronic state which is characterized by difficulty in attention control, increasing irritability, insomnia, nervousness, and disregard for danger. As the condition develops, the sufferer becomes chronically fatigued and stale and the safety ceiling is progressively lowered. The practical ways for meeting the dangers of anoxia are obviously through improved facilities for conveying and supplying oxygen, and by the training of flyers to be aware of the hazard and to become self-protective at the appropriate point, under operational conditions. To a substantial degree, this objective has been realized. But new operational demands pose new problems. It will be obvious that even with the best possible oxygen mask, and when using 100 per cent oxygen, a man who is flying at 41,000 feet cannot achieve a saturation of the blood much higher than 88 per cent. To be effective at this altitude or higher additional protective devices are requisite, and to meet this need pressure cabin planes have been designed. This and other problems to which scientific effort is now being devoted will no doubt during this war and afterward find their adequate solution as men continually fly higher and farther.

Sub-zero temperature, usually about —13° F. at 20,000 feet, is an ever-present hazard in the upper atmosphere, because of its pronounced psychological and physiological effects on the organism. The consequences of cold have been less studied than those of anoxia and deserve more investigation than has been given to them. An interesting research was reported by Armstrong [11] on the effects of increasing cold (50° F. to —40° F.) on 35 pilots who flew in open-cockpit planes. Exposure in the open cockpit is not dissimilar to the current experience of the waist gunner in a military plane. Flying personnel even when protected by heavy clothing begin, at 10° F. to suffer from the cold in terms of pain, muscular stiffness, and loss of acuity of touch. Mental distress and loss of morale develop progressively, indifference toward or even dislike for their duties appears, and finally a tendency either to panic or to stupor. Tactile efficiency essential for plane control decreases from 100 per cent at 50° F. to 68 per cent at 0° F., 33 per cent at —20° F. and finally to 13 per cent at —40° F. The need for most careful experimentation and development of flight clothing which will protect without interference to skilled action is obvious.

Besides anoxia and cold, high-altitude flying predisposes to acute bends-pain and to temporary or chronic fatigue. "Bends" or "caissons disease" has been long known as a distressing experience of divers who ascend from increased pressures so rapidly that the nitrogen absorbed in the body fluids and tissues, changing to a gaseous state, produces dangerous and painful symptoms. Now we know that flyers, too, may develop similar distress when they rise to elevations between 35,000 and 38,000 feet and remain there 45 minutes or longer. There are large individual differences in susceptibility and in the seriousness of the effects. It may be that slender, physically fit individuals are less susceptible than others. An early symptom is itching of the skin; later comes pain in the limbs, especially in the knees and in the shoulders, in both muscles and joints. Abdominal pain results from distention of gas in the abdomen; and a feeling of ill-defined discomfort is accompanied by sweating, chilling, and dizziness. Psychologists are at work analyzing and grading the intensity of the subjective effects. But the physiologists must be depended upon to devise measures and routines for avoiding these effects. Three procedures have been recommended. One involves pre-breathing of pure oxygen at ground level for 30 minutes or longer before ascent (Fig. 13), to get rid of a portion of the nitrogen in the organism. Another calls for strict avoidance of strenuous bodily effort at the altitude where the effects of bends tend to occur. The third is the preselection of the less susceptible personnel by repeated tests made in decompression chambers.

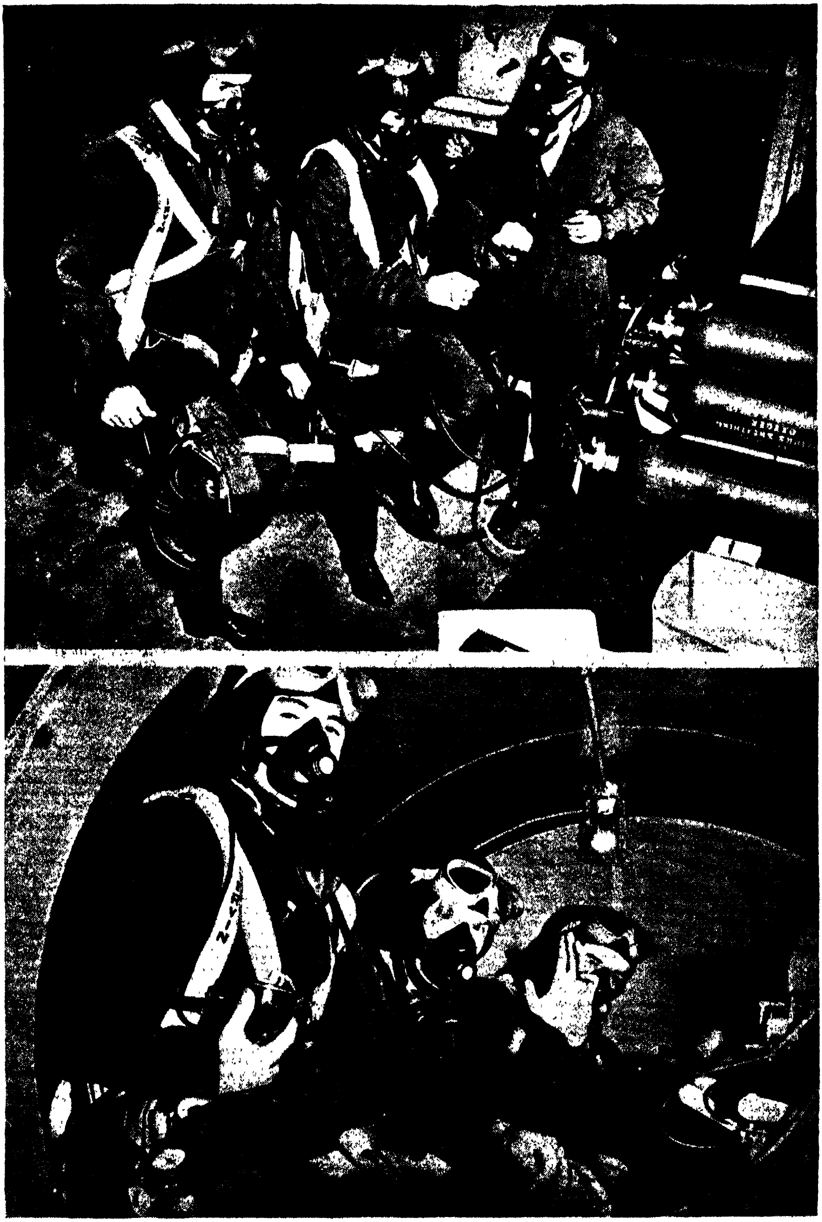


Fig. 13. Top, preparing for a high-altitude "flight" in the decompression chamber by exercising and breathing pure oxygen. This procedure rids the body of some of its absorbed nitrogen and protects against bends pains. Bottom, denitrogenized now, the trainees enter the decompression chamber. They continue to wear their masks even at sea level to avoid breathing in more nitrogen. (Courtesy of Boeing Aircraft Co., Seattle, Wash.)

There are numerous other problems relating to the protection and safety as well as to the efficiency of flying personnel which have received detailed scientific consideration. Many of these require the creation of specialized equipment. A partial list would include helmets, goggles, ear protectors, microphones, safety harness, emergency rations, medical kit, crash equipment, anti-g gear (to avoid blackout in dive bombing), signals to indicate the faithful operation of the oxygen system; and even the design and instrument layout of the cockpit requires consideration. All of this equipment is for use in certain specified ways and has to be mastered by flyers. Indoctrination and drill are required.

Some of the scientific work and development concerned with protection must be done outside of service or civilian laboratories. An example is the study of parachutes and parachute escape. A notable instance of a field investigation of this character (Fig. 14) is the well-planned experimental jump made by Lt. Col. W. R. Lovelace, acting chief of the U. S. Army Air Forces Aero Medical Laboratories, Wright Field, on June 24, 1943. He bailed out of a Fortress used as a "laboratory" plane at 40,200 feet over Boeing Field, Seattle. The jump was to determine the effect of extremely high altitudes on the human body and to measure the effectiveness of existing emergency parachute equipment. In place of a "delayed-opening" jump, as is commonly used in high-altitude jumps, the parachute was opened by a static cord, attached to a clamp in the bomb bay, immediately after he left the plane. He parachuted the entire distance to the ground. The temperature at the 40,200 feet elevation was —40° F. and he wore clothing appropriate for this, together with a helmet and goggles. He carried two oxygen bottles, each good for about 20 minutes' supply. He wore two parachutes, a standard Army and Navy 28-foot back-pack chute and a 24-foot chest-pack chute, the latter for emergency. The jump is reported to have been successful but the detailed scientific results, like the great majority lately secured in both service and civilian laboratories, have not been released to the scientific public.

FIG. 14. Lt. Col. W. R. Lovelace making ready for an experimental parachute jump from an altitude of 40,200 feet with the purpose of studying protection problems. (Courtesy of Boeing Aircraft Co., Seattle, Wash.)

FLYING FATIGUE

Flying sickness, or pilot fatigue, is a difficult condition to study objectively. However, scientific observers agree that it does actually occur as a serious occupational hazard. Apparently it is unlike ordinary bodily fatigue following excessive muscular work, and more like the fatigue of psychological depression. The total amount of muscular work performed by a pilot is usually relatively small, and the food-energy cost of this or anyone else's mental work is minimal in physiological terms. But psychologically the picture is quite different. The pilot's job is heavy in terms of continuous mental adaptations, responsibilities, and crucial judgments. He is often hampered in the ease of his manipulations by a load of distracting, but necessary, protective equipment. It is under requirements for prolonged flying with weather or military hazards present that typical flying fatigue appears as a serious condition. Learned response habits may remain essentially intact, but when the condition is serious parts of the habit system split off from the main pattern so that efficiency, usually attained through total habit integration, is greatly reduced. Another way to express this is to say that the timing integration goes to pieces. Every part and phase of the man's work requires more subjective checking.

In fatigue there may be an actual increase, rather than a lessening, in the work intensity. But this increase is at the cost of skill. With the fatigue comes a narrowing of attention. The pilot endeavors to do only what he feels he must do; what, in this ill-organized condition, seems absolutely essential. He may forget or ignore the temperature gauge, or the gasoline gauge, because he is completely absorbed in immediate plane manipulation. Like the mentally distracted or depressed person the fatigued pilot may tend to resort to random responses. It is as if he had forgotten the familiar resources available in his instruments and controls. In this condition he may feel, from his effort at concentration, that he is doing quite well. Elements of poor performance noted by him he tends to attribute to his plane or mechanisms within it. The importance is clear of continuing studies to determine critical points for lengths and types of flights and favorable treatment in case flying fatigue actually develops.

Military aviation is unavoidably hazardous. The flyer who professes "not the slightest fear" is bluffing. This is pardonable because it is one way, or at least an attempt, to bolster the sense of security. Another way is to convert the anxiety into something external and if possible objective and somewhat funny; hence the "gremlins." When visiting huts of the R.A.F. units in Britain in 1942 one could see on the walls drawings of "gremlins" and poems about them. The "gremlins" according to the best information available were first "observed" ("projected" would be a better word) in the photographic reconnaissance units. Here men in planes stripped of armor undertook lone missions in daytime over enemy territory. The pilot could scarcely think otherwise than that everything depended on the good functioning of his plane. At times, especially when tired, it was easy for the reconnaissance pilot to feel that something was bedeviling his engine, the gas feed line, or other vital features of the ship. The "gremlins" spontaneously took their grotesque shapes out of the enemy-charged, anxiety-loaded atmosphere.

In the psychology of the warrior this phenomenon is not new. He has always had to expose himself in ways that were inconsistent with his instincts for self-preservation and comfort. Otto van Veen has provided us with a sketch (Fig. 15), not from a soldier's hut of 1607, but from his book of that date illustrating selections from Horace. The reference is to Horace's *Odes,* Book III, page 1, lines 39-40, "nor does black care quit the brass bound galley and even takes her seat behind the horseman." "Black care" as a term is rather severe; it lacks the semi-humorous, nonchalant flavor. Probably the ancient soldier had a better word for it. In the P.R.U. huts of 1942 no one would ever confess to having seen a "gremlin." But "gremlins" had definitely become a part of the semi-humorous, danger-masking vocabulary that permitted these men engaged in military aviation to talk freely of things of which they had some fear.

MAINTENANCE OF OPERATIONAL EFFICIENCY IN AIRMEN

Psychological flyer upkeep is as necessary and important as mechanical plane servicing. Aircraft need refueling, rechecking, overhauling. Airmen require vigilant personal maintenance. Viewed from without, a group of planes flying in formation (Fig. 16) may look perfect as a military unit. But the view from within is of equal importance for assessing the fighting strength. How well are the airmen in the planes psychologically fortified against an attack which may suddenly appear from below the clouds or out of the clear sky above? To be operationally top notch flyers must be sustained in first-class physical and mental condition. This job has been traditionally the responsibility of the flight surgeon, a medical man of general training and specific Army or Navy experience. Many flight surgeons are gifted, as they need to be, with the combined traits that make successful and admired football coaches and insightful and effective deans or guidance specialists. When they have been trained as psychiatrists, so much the better.

Flight surgeons who understand the nature of the psychological strains involved in operational and combat flying put to good use their close association with their men and the opportunities offered by the periodic medical and physical examination check-up. This type of psychological-medical service has not been standardized to any great extent, and it is little taught as a medical specialty outside the two excellent schools: Randolph Field for the Army and Pensacola for the Navy. Psychiatrists in the service have in the past given greater emphasis and more attention to specific problems of psychoneurosis and psychosis, the weeding out of the unfit as a preselection measure, and at later stages the care of those who become clearly ill mentally. In the present war emergency many physicians, psychologists, and others are working on problems of morale, detection of flight fatigue especially in its early stages, and appropriate therapy for remedial benefit. Adjustment problems associated with

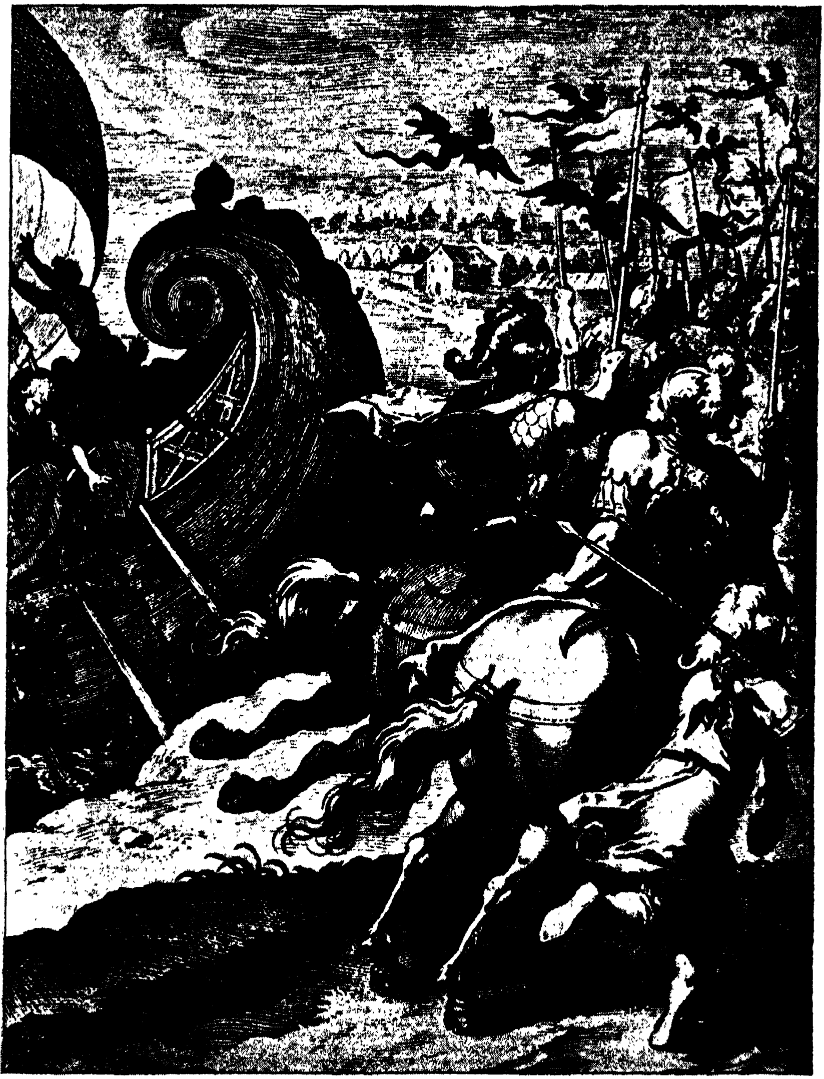


FIG. 15. An artist's representation of the ever-present "black care" devils which, according to Horace, beset the life of the soldier. These ancient devils appear to have their counterpart in the "gremlins" which are said to vex military flyers today. (From Otto van Veen, 2. Horatio Flacci Emblemata, Antwerp, 1607.)

acute fatigue can often be dealt with practically by well-planned direct or indirect methods. Chronic fatigue is less frequent, more insidious, and more taxing on the time and wisdom of the flight surgeon. The pilot may not recognize nor be willing to admit the accumulative emotional and mental effects that the strain of combat flying has brought about. He does not want to let his teammates down, he blames himself for his waning spirits, and he tries to force himself harder in order to recover. Therapeutic measures have to be tactfully brought to bear on such a flyer by his medical-psychological adviser even against the man's own first insistence that he needs no help. The situation demands understanding, skill, and knowledge of all his men by the flight surgeon. Fortunately not a few of the flight surgeons are themselves flyers. A pilot will take much more from an adviser who, he feels, identifies himself with his own personal and operational objectives. The flight surgeon needs to know the man's interests, abilities, ambitions, his personality, and indeed his total history: what has he been up against in earlier experience and in recent operations.

Through listening and watching rather than through questioning, the helpful medical officer can generally keep in touch. An operations program that allows for suitable diversification in work, rest, and recreation will produce few break-downs. The flight surgeon's appraisals can be used effectively by the commanding officer in making assignments. A reasonable man-power program within each operational group, with suitable flexibility to meet the human needs involved and with suitable stimulus to preserve morale at a maximum, generally allows a man to compensate for unavoidable fatigue and so to steer clear of its evil consequences. We need not here describe in detail the symptoms of chronic "flying stress." That it occurs relatively infrequently is an evidence of adequate medical and psychological care and of good operational organization. Both the acute and the chronic conditions involve psycho-physiological factors which are under continuous study.

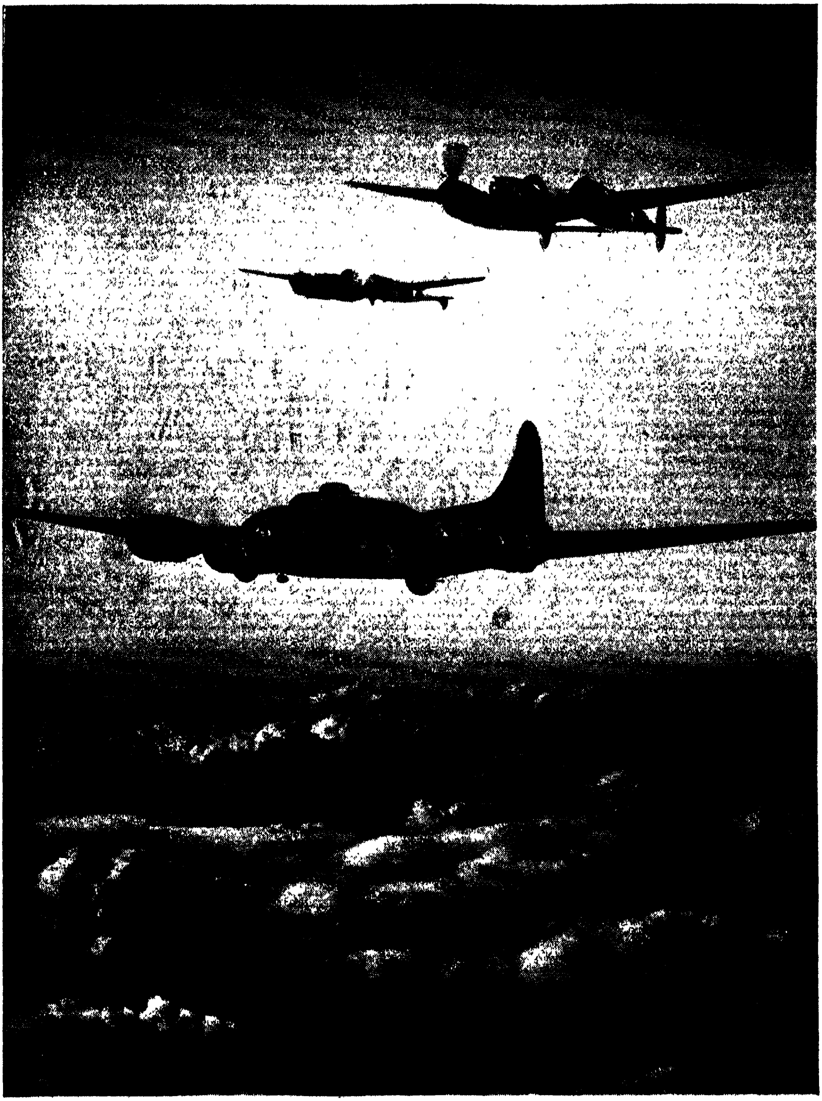


Fig. 16. Flying Fortress with its protective cover of Lightnings ready to engage in deadly team work. The striking power depends not on the planes alone but on the training, protection, and maintenance of the men who fight through them. (Composite photo, courtesy of the *Aeronautical Engineering Review.*)

The contribution of morale is highly important. Pilot stress is increased by frustration and lack of success, whether general or specific. The achievement of desirable and attainable goals raises morale and increases energy. One procedure that has been found to assist morale is the definition of duty in terms of hours of combat or number of missions completed. Such specific tasks can be specifically fulfilled and a positive score chalked up. Recognition of merit in terms of the heroism of routine service, either general or combat, is psychologically sound. Five combat flights deserve the air medal they receive, and give some lift of spirit which the recognition inspires. The men who fly are relatively little subject to moods, yet like all other human beings they experience positive and negative affect in their relation and attitude to people and events. When physical stress is great, psychological stress may result as a secondary consequence. The treatment consists in good medical care with sound psychological understanding. Morale building involves a continuous course of immunization to the chief psychological hazards, first in operational, and later in specific combat experience.

CONCLUSION

The study of man as an aviator has come to include an increasing series of typical flying and fighting characters. In each of his roles, as the psychologist analyzes them, man is viewed as being, not as playing, the character studied, in this case the aviator. This character differs from man as athlete or college student, as artist, journalist, businessman, or physician. Psychological study of personalities in the diverse roles they take shows that the observed characters represent true life-behavior patterns. They are not just players, i.e., players' masks and costumes to be put on or taken off at will. The man in the role is the functioning man, and in essentials of type each character group differs from every other one.

In a national and world situation unique in its urgent practical needs and presenting a vast human material for study, psychologists are well aware of their serious responsibilities and their inadequacies. Limitless studies are theoretically possible, but only those that can give useful results seem at the moment practically defensible. The scientific program of research must steer a middle course between too elaborate detail and too superficial generalizations. The investigator must return constantly to correct and to be guided by an accurately designed and carefully executed blueprint of definitely formulated hypotheses. Two questions are fundamental as points of reference: (1) Who is the ideal aviator, and how did he achieve this character? and (2) Which of the thousands of aspirants have the traits to become ideal aviators, and how can we, by training and with protection and care, increase their chances of growing most effectively and most quickly into this role?

Both questions have ultimately to be answered in terms of the integration of partial into total personality function and development, both typical and individual. The immediate fractioning of personality and its growth, into separate behavior samples, viewed in action, is the first task. This step is being taken deliberately and with forethought.

What results may be expected or hoped for from this great national research? First, the store of techniques and the sum of knowledge acquired through their use will benefit by the opportunity for large-scale validation or disproof of accepted method and results, by renovation of materials, and by the addition of new devices for measuring, teaching, protecting, and maintaining groups of human beings under special and currently developing conditions. A similar opportunity in World War I led to extensive job analysis, test construction, and measurement of specific traits and skills of men in service. Advance in the same lines and also in many others is now being made along a wide front by the concerted action of many individuals and groups of investigators. Second, orientation of study and measurements, not only to the problem and the specific trait but to the individual personality as a functioning whole is a contribution of psychology that serves as a focal point in many present studies. Advance in theory and in methodology may be hoped for from this psychoörganic point of view. Third, the coöperative nature of current research is bringing to light new approaches and new problems. Unbound by the historic prejudices and assumptions of psychologists, men in other fields, for example, physiology, biochemistry, biophysics, internal medicine, and psychiatry, can often suggest by their questions and demands wholly new aspects and needs, and sometimes also the solution to old resistant problems.

In the early periods of airplane development the problems of the aviator appeared relatively simple. The self-chosen, self-trained pilot had what it took. Modern engineering design and construction have shifted the balance. The machine can do more and requires more from the flyer than he can, unaided, do, stand, or give. The personality engineer is pushed to the utmost to keep up with the construction engineer. Machines and man-power will decide the present bitter struggle of nations. They will also largely establish the level of operations for the peacetime takeoff. Selection, training, protection, and care of human beings as workers and as citizens in a constructive, free world civilization are patterned in the current national research. In the longer scientific view the beginnings of new developments as well as the improvements of old methods may be seen in the present exploration into a new realm, as well as concentrated attacks on several of man's oldest enemies—preconception, ignorance, stupidity, and inertia. The synchronizing of man's active psychological skills with the products of his mechanical and technical laboratories and factories may ultimately prepare him to master the material forces in his world which at present threaten to master him.

For References see p. 305.

II

THE PHYSICAL STRUCTURE AND BIOLOGICAL ACTION OF NERVE CELLS

WITH SOME REFERENCES TO THE PROBLEM OF HUMAN FLIGHT

By DETLEV W. BRONK
Johnson Research Foundation
University of Pennsylvania

THE nature and action of nerve cells are of interest to all scientists who would understand the translation of physical causes into biological effects, for it is in these structures that animal organisms have their most sensitive contacts with the physical universe. Through them the activities of man are controlled by the forces of nature.

I

To the biologist, nerve cells have a unique significance as the elements which integrate an assemblage of cellular units into a living animal body. Without their integrative action the organism would disintegrate into cellular chaos.

The continued existence of a living cell requires an environment of limited physical and chemical characteristics; if the conditions change beyond those limits, the cell structure is destroyed. In the simplest forms of animal life the activity of a single cell is under the direct influence of external forces. With the development of more complex organisms, however, most of the cells are separated from the outer environment and surrounded by other cells and by the fluids of the body. In this way they have a more protected existence and are able to preserve their unique structure, as long as the conditions within the body remain fairly constant.

This constancy of the internal environment against the influence of external forces is accomplished by bringing all the cells of the body into a coördinated pattern of activity. Every cell is thus able to maintain its own existence by playing a specific role in the activity of the whole organism. This harmony of action is largely due to the remarkable capacity of the nervous system for translating a sensitive reaction to physical and chemical agents into a coördinated control of the elements of the body.

The nervous system is itself an organization of countless cellular units. Each one is of microscopic diameter, but many of them are long enough to extend uninterrupted from the brain or spinal cord to the surface of the body, where they are exposed to the action of environmental forces. Others reach out to control the muscles and the glands, and between these two types there are vast numbers of interconnecting cells, which give to the nervous system its powers of integration and coördination. Each one is a discrete unit, but at its terminals it is in close contact with other elements of the system. At these junctions one cell acts upon others and thus contributes its influence to widespread action.

External forces initiate actions within the body through the medium of those nerve cells which serve as sensory outposts. These actions, in turn, tend to prevent the domination of internal conditions by the outer forces. A fall in the temperature of the surroundings, for instance, is followed by a lowering of blood flow through the superficial layers of tissues, and this prevents too great a loss of heat from the interior or light rays from a dangerous object focused, on nerve endings in the eye, may cause, through devious nervous pathways, the muscular movements of escape.

Not all threats to the constancy of conditions within the body are thus thwarted by anticipatory actions. Many external influences break through the defenses of the organism and modify the state of the interior. When that occurs, the nervous system helps to restore normal conditions. The initial agents in this corrective action are nerve cells which have a special sensitivity to one or other of the critical factors in the internal environment. A sequence of events is thus initiated which modifies the activity of cells and organs throughout the body. This limits the internal effects of external causes.

In the course of evolution, organisms have developed which differ widely in respect to these relationships between animal life and its physical environment. The life of some forms depends upon unchanging surroundings; others possess the ability to cope with a wide range of physical and chemical forces. In varying degrees the activity of an animal's body is limited by the structure of its skeleton, the power of its muscles, the chemical composition of its body fluids, and especially by the manner in which its structure and actions are modified by the influence of external agents on the system of nerve cells.

For the human organism, these relationships are profoundly modified by the use of machines and by the creation of a controlled environment. As the physical scientist has acquired an understanding of the cause and effect of natural forces, he has been able to devise means for extending the powers that man had slowly acquired through organic evolution.

Many of the scientific discoveries which provide these human advantages required the use of instrumental aids to the senses. The direct action of radiant energy and mechanical forces on sensory nerve cells has revealed many natural phenomena. But it is by the telescope, the microscope, the balance, and by the instruments of modern science that we have gained much of our evidence of the universe in which we live. By the concentration or amplification of energy, or through the transformation of one form of energy into another which is more delicately perceived, these devices increase the sensitivity, range, and discrimination of the human senses. We have established a new sensory relationship to our environment through the medium of such instruments.

The train of sensory nervous events which are thus initiated may evoke intellectual processes which lead to an understanding of natural phenomena. That in turn has stimulated the invention of machines for increasing human power to deal with natural forces. Thus, man alone has acquired the power of adaptation to his environment with the initial, sensory aid of an instrument and with the final, motor assistance of a machine. By these means the scope and the security of his life have been greatly increased.

The biological suitability of a machine is, nevertheless, limited by the characteristics of the human organism, as well as by the ingenuity of the physicist and the engineer. The healthy progress of our technological civilization therefore requires that the human significance of machines should be recognized, lest great powers be inadequately controlled or the body be subjected to conditions which overtax its capacity for adaptation.

It is my purpose to describe briefly some of the characteristics of the nervous system which determine the relation of the human organism to its physical environment, and to show how these characteristics may influence the design of instruments and machines, or limit their usefulness.

II

Much of our knowledge about nervous action is deduced from a study of the effects it produces, such as muscular movements, glandular secretions, animal behavior, human sensation, and thought. The functions of the nervous system as a whole can, indeed, be studied in no other way. But we must seek the causes of those effects in the nerve cells of which the system is composed. Only there can one observe the sequence of events that begins with the sensitive reaction of a cell to a stimulus from the environment, and ends with far-reaching biological consequences.

When we say that a nerve cell is extremely sensitive to its environment, we imply that its molecular structure is readily altered by physical and chemical agents. It is this structural change and the associated electro-chemical reactions that constitute nerve action. Accordingly, the activity of nerve cells must be measured in terms of some change in their structure.

Despite the lability of the molecular and ionic organization of these cells, their gross form is relatively constant; it may, indeed, be fixed indefinitely by certain chemical substances. This has enabled those who dissect the animal body to trace the general outlines of the brain and spinal cord and the course of peripheral nerves. With the aid of a microscope and chemical stains the cellular elements have been revealed, as well as the general arrangement of cells into groups and tracts. It is also possible, by microscopic observation, to discover certain gross features which determine the mode of action of physical agents on sensory cells. The auditory receptors are thus seen to be connected with a membrane that is set in motion by sound waves in the surrounding air. And at a suitable distance behind the lens of the eye the visual sense cells can be detected, spread in a thin layer and containing a substance which undergoes reversible chemical change under the influence of light.

From each advance in microscopy or in staining techniques there has come additional information regarding cell structure and the electron microscope now promises to give us greater resolution. Scott and his associates [18] have already contrived to put this new tool to good use in this field by taking advantange of the fact that the emission of electrons from its thorium-coated cathode is increased by the presence of calcium. Accordingly, when a thin section of carefully dried nerve cells is placed on the heated cathode, the focused beam of electrons shows where the calcium is located. This biologically important ion is thus seen to be concentrated around the nucleus and at the surface of the cell, where it is in contact with the environment (Fig. 17). This is significant, for we shall see that calcium plays a major role in determining how great is the cellular effect of an external agent. The method of observation is open to the criticism that it deals with the remains of a cell that has been subjected to the violent action of high temperature and drying. That the distribution of calcium is consistent with other evidence is, however, reassuring and offers hope of future usefulness for the electron microscope in the study of nerve structure.

Other recently developed methods for studying the basic structure of nerve cells employ diffracted x-rays or polarized light (Fig. 18). By these means Schmitt and his associates [17] have analyzed the molecular arrangements within cells which have been kept in a normal condition. Such studies are accordingly of especial value for the relation of structure to function.

The chemical composition within the cell boundaries, outlined by these optical methods, is different from that of the surrounding medium. This is due to the unique molecular organization and content of the cell surface. In the case of certain nerve fibers the differentiation is, indeed, so marked that a fluid interior can be withdrawn, leaving an outer, sheathlike membrane. Chemical analyses then show that the ionic composition of the two parts is different; also certain enzymes which have an important role in nerve activity are found to be concentrated in the outer structure.

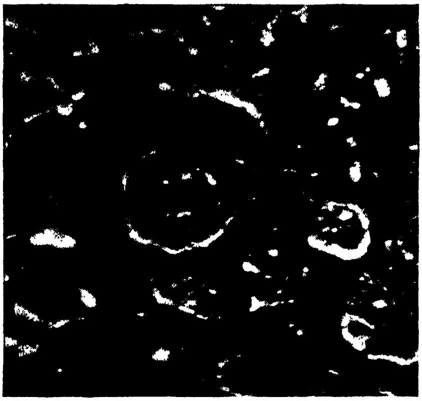


FIG. 17. The distribution of calcium in nerve cells photographed by means of an emission type electron microscope. Obtained by mounting a dried section of the nerves on the cathode. Boundary of a cell and its nucleus clearly visible in the center of the field. (From G. H. Scott.)

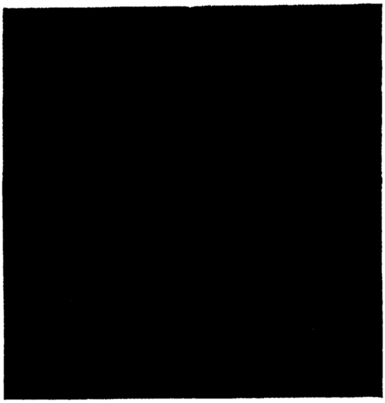


FIG. 18. Single nerve fiber from a crayfish photographed with the fiber in diagonal position between crossed Nicol prisms. The lighter portions indicate positive birefringence. (From R. S. Bear and F. O. Schmitt [4].)

A significant fact disclosed by these chemical analyses is the distribution of electrolytes within and without the cell. Potassium is present in much larger concentrations in the interior. Sodium is more concentrated in the surrounding fluid. And calcium, although more abundant in the cell as a whole, is largely concentrated in the surface structure. These are important elements in the organization of a nerve cell;

if the amounts in which they are present be altered by external influences, the characteristics of the cell are profoundly modified. That they do not freely diffuse to and from the outer fluid medium is due, in large part, to a quality of the surface membrane which tends to conserve the concentration of these essential ions.

This important characteristic of the bounding surface has been investigated by Cole and his colleagues [11], by measuring the resistance to the flow of electric current. Using certain nerve fibers that have a diameter as large as 0.5 mm., they were able to place one electrode in contact with the interior and a second on the outer surface. Thus they found that the resistance across the surface is high compared with longitudinal resistance along the fiber. These measurements support a conception of the cell surface as a structure that is relatively impermeable to ions and is more highly organized than the interior.

Because of these properties of the cell boundary there is a gradient of electrical potential across the surface. This important characteristic has long been known, but present development of physical and biological techniques for dealing with the fiber of a single nerve cell has recently enabled Hodgkin [14] to show that the difference in potential between an electrode placed in contact with the inside of the surface and another on the outside may be as much as 90 millivolts. The magnitude of this potential gradient depends upon the condition of the boundary, and we shall see that it is, accordingly, a sensitive measure of the state of the cell.

III

At this point I should emphasize the sensitivity of the surface of a nerve cell to the influence of external agents. It is readily altered by mechanical forces, specific chemicals, or by an electric current.

The surface effects which they produce are various. The general permeability may be so much increased by some external conditions, that the inner substance of the cell diffuses freely into the surrounding fluids, and the cell is destroyed.

Other agents cause a reversible change of structure from which there is rapid recovery, when the stimulus of the external action is removed. Or the stability of the surface may be increased, so that the stimulating action of other agents is inhibited. Whatever the nature of the change, it alters the relations of the cell to its surroundings.

To study these reactions of a cell to environmental forces, it is necessary to employ a stimulus which can be regulated precisely. For this purpose a brief electric current has been favored by the biologist because the strength and duration of its action are easily controlled. It also has the virtue of being a mode of excitation which plays an important role in the natural propagation of the nerve impulse. A change in the concentration of some chemical component of the environment is, however, a more selective means of excitation, and it reveals the specific role of that substance in the activity of the cell.

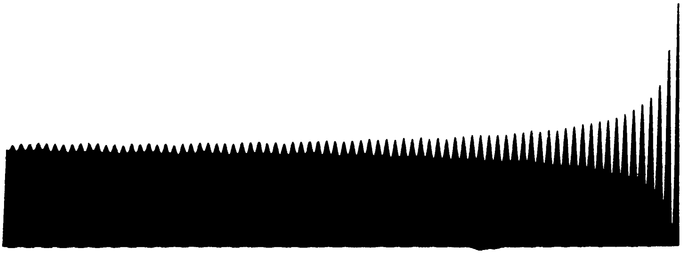


Fig. 19. Rhythmic oscillations of potential in the region of a nerve fiber from which calcium has been removed. Fiber is the giant axon of the squid [6].

The partial removal of calcium, for instance, makes a nerve more susceptible to the influence of other external agents, and from this we may conclude that its presence gives stability to the cellular structure. This fundamental property may be demonstrated by recording the potential of a point on a nerve fiber from which calcium is removed [6]. Normally this point has a constant potential, equal to that of other points along the cell. But as the calcium is withdrawn from the region under observation the potential changes, not steadily but in an oscillatory manner (Fig. 19). There is thus revealed a remarkable capacity of the nerve structure for periodic change and restoration. The magnitude of each cyclic alteration is governed to some extent by the amount of calcium which remains, or as we might say, by the stability of the cell. The frequency with which the alterations recur is, however, relatively constant. So we must assume that the characteristics of the cell structure which govern the periodicity are not easily modified.

If the degree of change in any cycle is sufficiently great, a second noteworthy phenomenon occurs. There then appears an alteration in the potential of points along the fiber far removed from the site of chemical action. A wave of change that can be detected by the electrical sign spreads along the surface, modifying as it goes many of the physical and chemical characteristics of the cell. Each altered section is in turn the stimulus for that ahead. The rhythmic activity of a local region, induced by an environmental stimulus, thus controls the behavior of the entire cell.

This rhythmic instability is also evident in the reaction of a sense cell to a constant stimulus. It is difficult to investigate that portion of the cell in which the action originates, but means have been devised for recording, in its extended fiber, the brief variations of electric charge which accompany the complex sequence of events that comprise a nerve impulse. We thus find that the action of a steady stimulus—light or mechanical force or chemical substance—is transformed by the cell into a train of recurring impulses (Fig. 20). The physical and chemical conditions produced by these impulses in the brain and spinal cord act as a stimulus to other nerve cells, and they in turn discharge a sequence of repetitive impulses to the muscles and the glands. A rhythmic change of structure, propagated to every part of the cell, is the natural mode of action of the units of the nervous system.

The discovery of this fact has been made possible during the past twenty years by the development of methods for recording the electrical signs of activity in single nerve cells. Earlier knowledge of the structural organization of the nervous system had to wait upon optical means for identifying the cellular elements of which the system is composed. Similarly, our present understanding of the basic events which constitute nervous action remained obscure until means were devised for observing the behavior of the individual cellular units, while in their normal position within the body. In making this possible, the electronic amplifier and methods of micro-dissection have resolved the basic events of nervous action, just as the microscope and staining techniques had previously resolved the structural units.

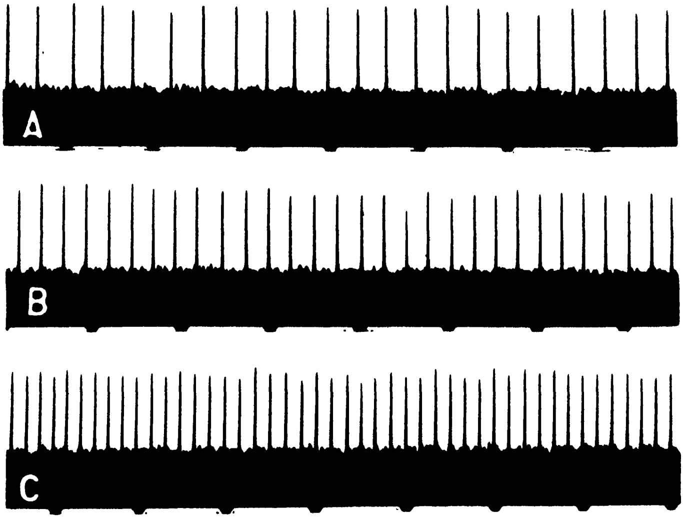


FIG. 20. Electrical signs of nerve impulses conducted along a sensory nerve fiber. Impulses initiated by constant pressure on a pressure-sensitive ending. The pressure is successively greater in A, B, and C. Time signal at bottom of records 1/5 sec. [10].

The transient variation of surface potential that identifies the passage of a nerve impulse lasts but a few thousandths of a second and has a recorded magnitude that may be no more than twenty-five millionths of a volt (Fig. 21). These conditions make exacting demands of the recording system, but vacuum tube amplifiers and cathode ray oscillographs have admirably satisfied the requirements. When combined with biological techniques for placing electrodes on single, intact nerve cells in experimental animals, they have enabled us to probe the events which occur within the sense organs, the brain and the spinal cord. These investigations which have been largely carried out by Adrian and his associates at Cambridge [2], by Gasser [13], and Lorente de No [15] at the Rockefeller Institute and by my colleagues at the University of Pennsylvania [9] have resolved into regular trains of cellular events many complex phenomena of the nervous system. In terms of these unitary actions, it is now possible to comprehend how the integrated nervous control of the organism is accomplished and how the organism is thus related to its physical surroundings.

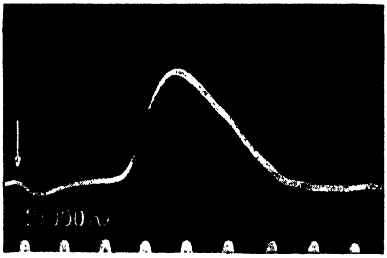


FIG. 21. The variation of electric potential at a point on a nerve fiber, measured with reference to an inactive point. The instant of stimulation is indicated by the arrow. Time signal .0001 sec.

If the nervous system is to serve as a medium for controlling the organism, in accordance with the conditions of the environment, its activity must in some way be modified by the quality and the intensity of the stimulus. I may say at once that the magnitude of nerve impulses is not such a measure; the altered character of the cell in the region where the action originates is merely a trigger for the self-propagated disturbance. The nature and degree of surface change at any point along the nerve fiber are determined only by the condition of that region. The magnitude of the conducted nerve impulse is, accordingly, independent of the nature of an adequate stimulus.

The initiation of this invariable, propagated disturbance does, however, depend upon the relation of the cell structure to the external agent, and that varies from cell to cell and from moment to moment. Furthermore, cells differ in regard to their sensitivity to different types of stimuli, either because of accessory factors such as the membranes of the ear or because of innate structural characteristics. This will determine which type of cell is brought into action, and the central destination of their fibers will, in turn, decide which portions of the brain receive the sensory messages. Sense cells of the same type differ somewhat in their sensitivity to the same stimulating agent. Because of this a more intense stimulus brings more units into action, with a correspondingly wider distribution of impulses within the nervous system. How these factors ultimately account for quality and intensity of sensation is beyond the limits of this discussion.

In the activity of each individual nerve cell there is a second measure of the intensity of the stimulus. This is the frequency at which the cell discharges impulses along its fiber (Fig. 20). If the stimulus be intense, the cycle of cellular change recurs frequently. Conversely, a train of impulses separated by long intervals is a message from a sensory ending exposed to a relatively weak stimulus. Through these variations in the frequency of nerve impulses going into the brain and spinal cord, the intensity of environmental stimulus grades the activity of those centers. Each succeeding cell along the nervous pathway in turn exerts a variable influence in this same manner (Fig. 22).

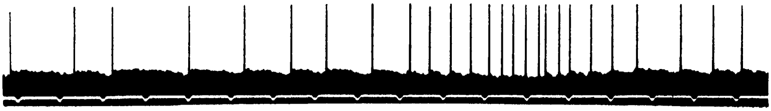


FIG. 22. Electrical signs of nerve impulses conducted over a single nerve fiber from a motor nerve cell to a muscle. Increased muscular contraction was caused by the more frequent activity shown to the right of the record. Time signal at bottom of record ⅕ sec.

In many types of sense cells the frequency of action follows changes in the strength of the stimulus with remarkable fidelity. The frequency is, accordingly, a precise measure of those changes. It is far less reliable as a measure of the actual intensity of the stimulus. Slight differences between cells cause them to respond at different rates. Furthermore, the character of each cell is progressively altered under the influence of the stimulus; and in this state of adaptation the cell reacts less frequently (Fig. 23). Because of these properties, nerve cells provide no definite index, constant from moment to moment, of the strength of forces in the environment.

On the other hand, the frequency of action of most types of sense cells follows changes in the strength of the stimulus with remarkable fidelity. A change of frequency is accordingly an accurate measure of alterations in the environment.

There have been many speculations as to the cellular mechanism through which variations in the stimulus modulate the frequency of the nerve message. They are all based on the conception of a nerve cell as an unstable structure which changes under the influence of the stimulus and thus initiates a nerve impulse. Recovery then follows, and not until the nerve has been somewhat reorganized is the continuing stimulus able to set up another impulse, and so another and another. It is assumed that if the stimulus be more intense, recovery need not proceed so far in each cycle before another impulse can start on its way. Thus the frequency is increased.

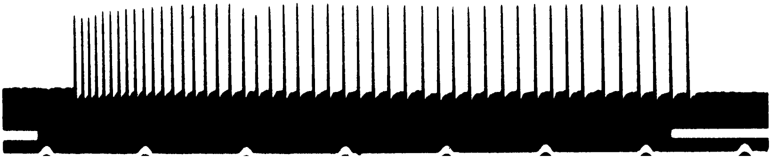


FIG. 23. The discharge of nerve impulses from a photoreceptor stimulated by light of constant intensity. The decline in the frequency of action, toward the right of the record, is due to the process of adaptation. Time signal ⅕ sec. (From H. K. Hartline.)

IV

The existence of a nerve cell as a complex organization of molecules and ions requires a continual expenditure of energy derived from chemical reactions. Carbon dioxide and other products of combustion are liberated, the evolution of small amounts of heat can be measured by sensitive thermo-electric devices, and oxygen is steadily consumed. If the cell is deprived of oxygen, it soon loses its power of rhythmic response to external stimuli.

The amount of oxygen required to maintain the cellular structure depends upon the state of the cell. This may be demonstrated by placing nerve fibers, removed from the body, in a closed vessel. The medium surrounding the nerves can thus be controlled, and the decrease of oxygen in the container is a measure of the amount consumed. If the structure be made more stable, as by the application of a narcotic, oxygen is used less rapidly. Conversely, it is used more rapidly in maintaining the less stable structure that results from the withdrawal of calcium. For any set of environmental conditions there is a corresponding cellular state, with a characteristic rate of oxygen consumption.

It is more difficult to measure the oxygen consumed by nerve cells as they carry on their normal functions in the body. Some information has been obtained by comparing the amount of oxygen in the blood entering and leaving the brain, but the results are not specific with regard to either place or short periods of time. This is an important limitation in view of the varying activity of the cells. A method recently developed by two of my colleagues, Davies and Brink, [12], has the advantages of both quickness and spatial discrimination.

In this case, the measurement of oxygen concentration in the environment of the cells is made by an adaptation of the polarographic method well known to chemists. A difference of potential is maintained between two electrodes. One, which is nonpolarizable, is applied to some remote part of the body; the other, sheathed in glass down to a metallic tip of microscopic dimensions, is placed in the immediate environment of the nerve cell. With an appropriate difference of potential, the electric current flowing through the circuit depends upon the rate of reduction of oxygen. The measured current is accordingly proportional to the concentration or tension of oxygen near the active electrode. By placing such an electrode in the fluid medium that surrounds one or a group of nerve cells, it is possible to measure the concentration of available oxygen.

To determine the rate at which oxygen is consumed, from the measurement of its concentration, it is necessary to control the supply. Useful for this purpose are small, compact groups of nerve cells, known as sympathetic ganglia. For our present purpose, they possess the advantage of a local circulation which can be regulated, and there is a set of incoming nerve fibers over which impulses can be sent to activate the cells.

When the flow of blood to one of these ganglia is arrested, there is a steady fall of current passing through an oxygen electrode that has been thrust into the midst of the cells. The rate at which it declines is proportional to the oxygen consumed by the cells for the maintenance of their resting structure (Fig. 24).

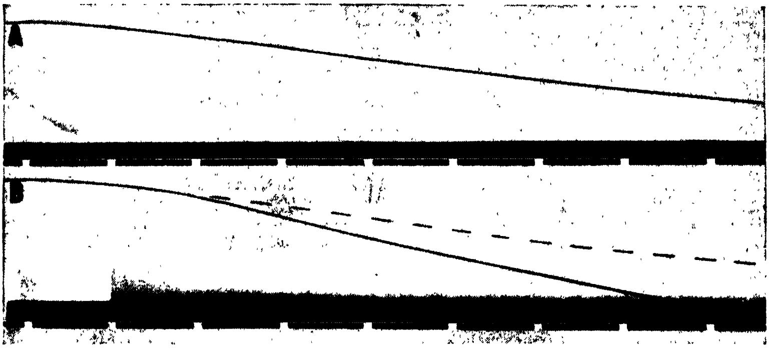


FIG. 24. The decrease in oxygen concentration within a group of nerve cells (sympathetic ganglion), measured by means of an oxygen electrode. Circulation stopped at beginning of both records. In A the cells were at rest; in B they were stimulated rhythmically. Upper line, oxygen concentration (current flow). Lower black edge in B, record of impulses discharged from cells. Broken line at bottom, time signal 10 sec.

Rhythmic activity of the ganglion cells may be induced by electric stimulation of the entering nerve fibers. When this is done, the rate of oxygen consumption is increased in proportion to the frequency of nerve action. This increase provides the energy necessary to reorganize the structure of the cells, following each cycle of cellular change. The recovery is a process of considerable duration. Measurements have shown that heat production rises during a period of activity, and only slowly decreases to the resting level during a period of some minutes. Likewise the increased oxygen consumption long outlasts the train of rhythmic activity. Because of this, the concentration of oxygen in the environment of the cells is decreased. How far this falls, and for how long, depends upon the flow of blood which carries oxygen to the cells.

The dependence of nervous activity on the circulating blood is especially significant in the brain, where the cells have a high rate of metabolism. In experimental animals, or in the course of human operations, we have applied the oxygen electrode to discrete points of the cerebral cortex, located through a microscope (Fig. 25). It is thus possible to study the spatial and temporal course of the oxygen exchange between blood and the cells. As the electrode is moved from the surface of a small artery, there is a steep decline in oxygen concentration until another vessel is approached where it again rises (Fig. 26). Whereas the oxygen tension of blood in the arteries is 100 mm. Hg., and 40 mm. Hg. in the vein, present measurements indicate that it may be less than 10 or 15 mm. Hg. at points .1 mm. distant from a vessel. This steep gradient of concentration is an important factor in the diffusion of oxygen from the vessels to the cells. Their avid consumption keeps the concentration low, and this is turn favors a rapid transfer from the blood.

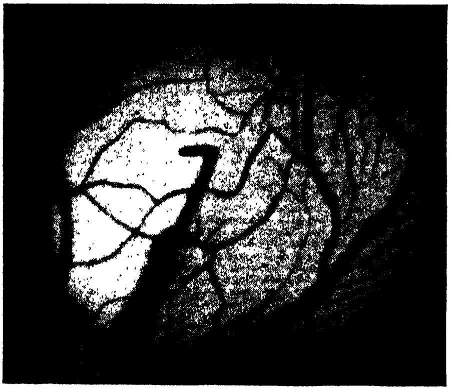


FIG. 25. Oxygen electrode in position on the surface of the cerebral cortex, in the midst of small blood vessels.

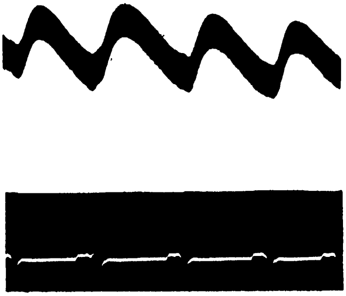


FIG. 27. Upper record: periodic variations of oxygen concentration at a point on the cerebral cortex .025 mm. from a small artery, following each respiratory cycle. Lower record indicates respiration.

The amount of oxygen available for diffusion depends upon the rate at which it is delivered by the blood. This is determined by the amount supplied to the blood in the lungs and by the flow of blood through the brain. The striking influence of the first factor is illustrated by our observation that the oxygen concentration rises and falls with each respiratory cycle, if the respiration is not too fast (Fig. 27). This shows that even a brief decrease in the oxygen taken by the blood from the lungs causes a deficiency in the brain. In the same way, a slowed flow of blood is promptly followed by a decrease of oxygen concentration among the cells, and a short stoppage leaves them altogether destitute. Obviously, there is in the brain a very short reserve of this material so necessary for the continued activity of nerve cells.

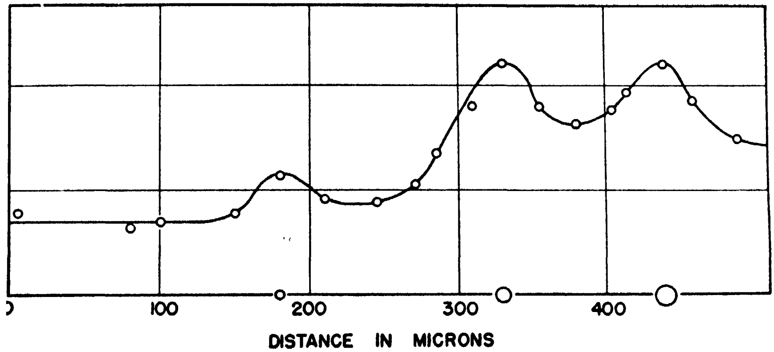


FIG. 26. Oxygen concentration at a series of points on the cerebral cortex, measured by the oxygen electrode. Positions of a vein and two small arteries are indicated on the axis of abscissae.

V

At the beginning of this discussion I suggested that the use of a machine for increasing some of our physiological powers may be restricted by other physiological limitations. This intimate relationship between machines and the human organism is dramatically illustrated by the important, if obscure, influence that the properties of nerve cells have had on the progress of aviation.

One such limitation is their oxygen requirement which we have just considered. In 1803, only twenty years after the first balloon ascent, Robertson, a French physicist, described his state at 22,000 feet as that of indifference.

"There," said he, "the physicist is no longer sensitive to the glory and passion of discoveries; it is only by the aid of a little fortifying wine that he succeeds in finding intervals of mental clarity and power." And later, in 1862, Glaisher, the meteorologist, and Coxwell, his balloon engineer, ascended to 29,000 feet where "the feeling of indifference" passed into stupor, and Glaisher lost consciousness. Death would probably have followed had not Coxwell, paralyzed though he was, "seized the valve cord between his teeth and pulled it two or three times by shaking his head violently." Man's new machine-aided powers had carried him into an environment unsuitable for life.

There was much debate as to the cause of these ill effects. Some thought it was the lowered barometric pressure; others believed it to be "the too small quantity of oxygen contained in the gas in which they floated." Finally Paul Bert, the French physiologist, put the question to the test of experiment. This he did by placing men in a chamber, where he subjected them to one or the other of the two conditions. The harmful effects of high altitude were thus shown to be due to the low concentration of oxygen.

Against this hazard, the body has certain protective mechanisms observed by even the earliest aeronauts. Referring to a flight he made in 1804, Gay Laussac reported that his pulse and respiratory rate were much accelerated. In the light of our present knowledge, we know that this was due to the fact that there are in the walls of certain blood vessels, leading to the brain, nerve cells which are sensitive to a low concentration of oxygen. When they are stimulated, they start a train of events in the nervous system, which increases the rate of respiration and augments the flow of blood through the brain. But if the ascent continues, the concentrations of oxygen in the atmosphere and in the blood ultimately become so low that the mechanisms of compensation are inadequate. At those altitudes insufficient oxygen reaches the cells of the brain, and they lose their power of rhythmic action.

"I certainly shall not take it upon myself to set the limitations of human activity and indicate the point, if it exists, where nature tells the aeronaut: 'You shall go no further,'" said Glaisher eighty years ago. Since then, human ingenuity has devised aircraft capable of operating more than ten miles above the earth. But man's physiological limitations are now as they were in the pioneer days of human flight. Modern aircraft have given man new powers he cannot use—unless he has the assistance of other instrumental aids.

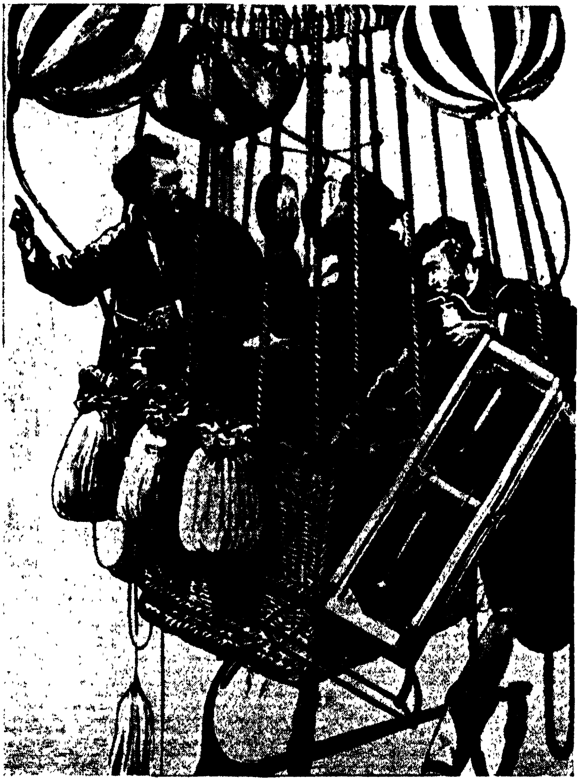


FIG. 28. The first oxygen supply system for high-altitude flight. Tubes lead from three bags containing oxygen [5].

The nature of the assistance required by man at high altitudes was first suggested by Paul Bert to the balloonists Crocé-Spinelli, Sivel, and Tissandier [5]. He constructed for them bags of oxygen, with tubes through which they could breathe this gas necessary to supplement the deficiency in the rarer atmosphere (Fig. 28). Fortified in this way, they made the classic ascent of April 15, 1875. They reached an altitude of 26,000 feet, but only Tissandier survived to recount the adventure. In order to conserve the supply of oxygen, they had delayed its use until mental confusion made them unaware of their need.

Today, military operations require frequent flights above the altitude reached by Tissandier and his associates. Careful instruction in the oxygen requirement of nerve cells has done much to warn our air crews of the insidious dangers of high altitudes, but the necessity for conserving the oxygen supply still remains. Every additional 100 pounds of oxygen tanks reduces the gas or bomb load by the same amount, so that it becomes necessary to reduce the weight of oxygen containers to the minimum that is compatible with human safety and efficiency. This has required the design of equipment which supplies just enough oxygen to the crew of a plane for their individual needs at any altitude.

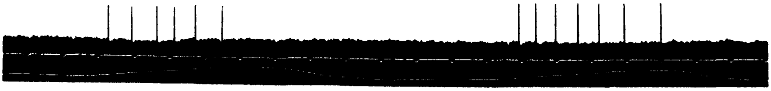


FIG. 29. Periodic trains of rhythmically recurring impulses in a nerve fiber which activates a muscle of inspiration. Below is the time signal (1/5 sec.) and a line which as it moves upward indicates inspiration [8].

Within the brain there are nerve cells whose activity is a measure of that need. They are the cells which periodically discharge trains of rhythmic impulses to the muscles of inspiration (Fig. 29). The intervals between the trains and the frequency of impulses within each group are largely determined by the concentration of oxygen and carbon dioxide in the blood. As each impulse arrives at a group of muscle fibers the fibers contract and relax, to contract again on the arrival of the next impulse. If the impulses recur more frequently, there is less time for muscular relaxation, the tension developed is greater, and the depth of inspiration is increased. The rate of respiration is, in turn, regulated by the intervals between the groups of impulses.

To insure an adequate supply of oxygen in flight, the face of an airman is covered with a mask which is connected to the storage tanks through a valve that regulates the flow of oxygen. Each inspiration lowers the pressure within the mask and thus draws in a diaphragm mounted in the regulating valve. The movement of this diaphragm opens a port by an amount determined by the depth of inspiration (Fig. 30). Thus the respiratory nerve cells operate the valve through the medium of the inspiratory muscles. In this way, the supply of oxygen is controlled by the needs of the organism rather than by atmospheric conditions. If too little oxygen is present in the body, the activity of the respiratory nerve cells is increased, the valve is thus caused to deliver more oxygen, and so the internal environment is maintained constant. Through the aid of a secondary machine the airman has overcome his inability to fly to high altitudes in modern aircraft.

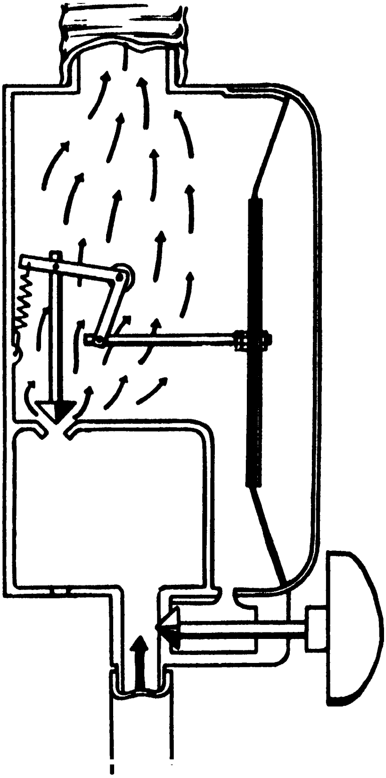


Fig. 30. Regulator valve which supplies oxygen in accordance with the depth and frequency of inspiration. Lead is to storage tank at the bottom; to face mask at the top.

*

Two of the tactical requirements of fighter aircraft and of dive bombers are great speed and high maneuverability. Engineers and metallurgists worked for years to develop such planes, capable of withstanding the centrifugal forces of high-speed turns. The machines they have produced now place new stresses on the human body.

A normal heart and circulatory system will deliver enough blood to the brain cells when the body is erect or recumbent, and will meet the needs during sudden changes of posture. This is partly due to the action of pressure-sensitive nerve cells located in the walls of certain blood vessels leading to the brain. With a constant pressure of blood in the vessels, the cells are stimulated to discharge a continuing train of impulses at a constant frequency. But every change of pressure causes the frequency to vary (Fig. 31). In this way the impulses regulate the activity of other nerve cells which control the rate of the heart beat and the diameter of blood vessels throughout the body. If the pressure forcing blood through the arteries to the brain decreases, the rhythmic activity of the pressure receptors declines. This causes an acceleration of the heart beat and a constriction of blood vessels in the lower parts of the body. Thus the flow of blood to the brain is increased again. Such a sequence of events tends to maintain an adequate concentration of oxygen in the brain, despite the redistribution of blood that accompanies changes in position.

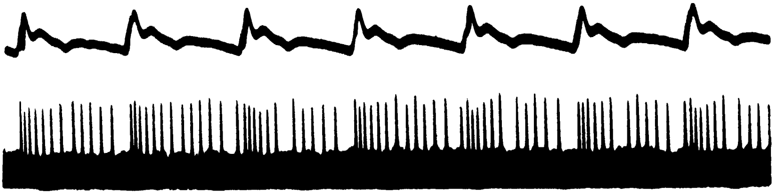


FIG. 31. Below, sensory impulses from a pressure-sensitive nerve ending in the wall of an artery, showing variations in frequency of nerve action with changes of pressure. *Above*, manometric record of pressure within the artery, varying with each heart beat [7].

This compensatory mechanism is, however, inadequate when the blood is driven toward the feet by a strong centrifugal force. During a high-speed turn, or in the sudden "pull-out" from a power-dive, the weight of blood may be increased five or ten times. Accordingly, it does not flow to the brain at a rate sufficient to supply the nerve cells with their required oxygen, nor are the nervous mechanisms of compensation able to correct the situation. The cells involved in the visual process first lose their capacity for rhythmic action, and as the oxygen deficit becomes more general, loss of consciousness follows. If airmen are to use their machine-given power of maneuvering at high speeds, it will be necessary to employ a device that will aid the heart to deliver oxygen to the brain, by preventing the accumulation of too much blood in the lower parts of the body.

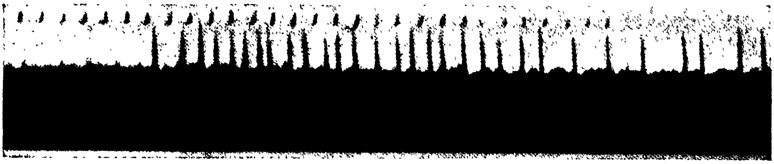


FIG. 32. Nerve impulses from a sensory nerve cell stimulated by the force of gravity during tilt of head [3].

The use of artificial wings and powerful motors has enabled men to escape from their earth-bound existence. We have thus gained freedom from the restriction of gravitational forces. But the full use of that freedom has required some new sensory adaptations to our environment, for the gravitational attraction of the earth plays an important part in determining the position of the animal organism in space.

The direct action of this force is to pull all parts of the body to the earth. But with the development of a bony skeleton and a system of muscles, animals gained the ability to stand erect and move about by raising their limbs against this force. In order that this may be done to good effect, the muscular contractions must be controlled in accordance with the orientation of the body relative to the surface of the earth. Vision plays some part in this control, and there are other types of sensory nerve cells that are stimulated by the tension of the muscles, by the rotation of the head, or by the force of gravity.

Movement and position are reported to the brain by the rhythmic action of these cells. Sudden rotation causes the discharge of a brief train of impulses from one type of sensory nerve ending. Cells of another type respond to the steady force of gravitational attraction, with a frequency of action that depends upon the orientation of the head to the direction of the force (Fig. 32). Any change of position is thus reported in the sensory message. These impulses, and those from the eyes and muscles, play an important role in regulating the activity of the muscles that control the posture and the movement of the body.

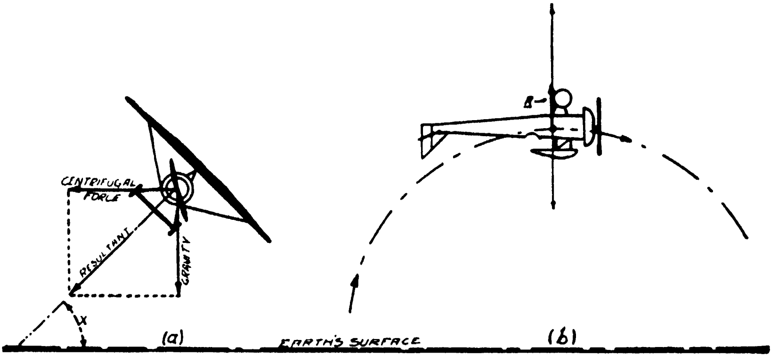


Fig. 33. Forces acting on a flyer in aerial maneuvers.

If a pilot is to operate his aircraft safely and effectively, he requires some contact with the earth through the medium of these sensory nerve cells. Photoreceptors in the eye provide him with such a visual reference to the land or sea. And in straight flight the gravity receptors are stimulated only by the force of the earth's attraction, so they provide an additional sense of orientation. But when the course is altered in a steeply-banked turn or in a loop there is the additional action of a centrifugal force (Fig. 33). Because the nerve cells are unable to distinguish between these two forces, their response is determined only by the magnitude and direction of the resultant force. Accordingly, in a turn the surface of the earth seems to be inclined to its true position. At the top of a loop the pilot has the impression that the earth has changed position with the sky. Only with proper training do the visual messages dominate the mental impressions and correct the sensory consequences of the unnatural stimulation of the gravity and rotation receptors.

When a view of the earth is obscured by clouds or darkness the sensory impressions remain distorted by the forces exerted by the machine. This loss of a true sense of spatial orientation then makes accurate flight impossible.

Such was the situation in the nineteen-twenties. A sudden change to cloudy weather, or a flight accidentally prolonged into the night, deprived the pilot of his visual contacts with the earth, and sensory confusion followed. Prolonged aerial operations were therefore uncertain and unsafe.

Among the flyers who rebelled against this human limitation on the use of aircraft was Major William Ocker of the United States Army Air Corps. Rebellion led to a study of his sensory reactions during sudden movements in the dark, and then one day while taking a routine medical examination he was subjected to a test in which he was slowly rotated while his eyes were closed. Under these conditions, he and the other pilots were unable to sense the true direction of their rotation. Ocker left this examination with the belief that finally he had found in the mechanism of his sensory nerve cells the reason for the difficulties of blind flight [10]. Soon afterwards he returned to the flight surgeon and requested that he be permitted to repeat the test with his head in a closed box, rather than with closed eyes. The physician agreed, and now Major Ocker reported each turn correctly. He had mounted in the box a turn indicator which he watched during his rotation.

It was more difficult to convince others of the correctness of an instrumental sign that belied the messages from the sense organs of posture, but he slowly proved the possibility of what has come to be called "blind flying." But it is not blind flight. The bank-and-turn indicator, the artificial horizon, the sensitive altimeter, the rate-of-climb indicator, and radar have increased man's power to use his aerial machines. They have done this by aiding the visual nerve cells to maintain contact between the organism and its environment. Only in this way can other nerve cells control the machine to human advantage.

I have cited these examples of the role of nervous mechanisms in the development of aviation in order to emphasize the biological basis of instruments and machines. If machines are to meet the needs of man, they must be designed to satisfy the biological requirements of the user. Only then can they fulfill their proper function, which is to increase the natural powers of the organism for dealing with environmental forces.

For References see p. 308.

III

ENERGY AND VISION

By SELIG HECHT

Columbia University

THE subject of this essay is specific and circumscribed, even though its title seems vague and general. It is concerned with a simple question: how much energy do we need to see light? In order to answer this question adequately, we need help from several sciences. We shall require both physics and chemistry, some biology, and even a dash of mathematics for flavor.

MEANING OF VISUAL THRESHOLD

How much energy do we need to produce a sensation of light? The first person to ask this question was Langley [18] in 1889, and the reason that he could do so was that he was the first person who could answer it. Eight years before he had invented the bolometer, an instrument for measuring radiant energy. A bolometer is essentially a blackened piece of wire whose electrical resistance varies with temperature. The blackened wire absorbs the light which shines on it, its temperature rises, and its changed resistance is then measured by appropriate means.

With this instrument Langley found that he could see a light when it deposited 3×10^{-9} erg on his eye. It may be recalled that at sea level it takes about 1,000 ergs to lift one gram one centimeter against gravity. One erg is therefore a thousandth of this energy, and in order to see we need only three-billionths of an erg.

A datum of this kind is impressive by virtue of its size as compared to man. However, an erg is an arbitrary unit, and by choosing some other unit, whose value for example would equal a micro-micro-erg, the amount of energy might be given as 3,000 such units and would surely not be nearly so impressive.

Langley's datum achieved interest in an absolute way with the advent of Planck's Quantum Theory in 1900. According to Planck, radiant energy is emitted and absorbed in discrete packets, which act as indivisible units. Each unit of light contains a quantum of energy. One cannot have half a quantum of energy, and this at once sets an absolute unit in terms of which one may express the amount of energy necessary for seeing.

For the light used by Langley, each quantum has a value of 3.6×10^{-12} erg. Therefore, the 3×10^{-9} erg found by Langley represents somewhat less than 1,000 quanta; and it is this number of absolute energy units which is required to start that complicated train of events beginning at the eye and ending in the cortex of the brain which results in our seeing light.

In 1911 Langley's measurement achieved significance a second time when Einstein formulated what has since become known as the Einstein Photochemical Equivalence Law. This states that for primary photochemical reactions each quantum of energy is absorbed by a single molecule which is then changed chemically by it. In short, one quantum absorbed means one molecule changed. Langley's measurement then says that for us to see a flash of light about 1,000 molecules of light-sensitive material in the eye are changed chemically by the light. This is really an astonishingly small chemical change.

We know now that Langley's measurements are wrong. But they are wrong in the right direction. The real value is even smaller than what he found; it is smaller by a factor of 10. This is mainly because Langley worked before the physiology of vision was known sufficiently well to enable him to use the eye most effectively. However, this is of minor importance compared to the establishment of an approximate order of magnitude. We have the fact that a biological process which involves such a high level of organization as a conscious act of seeing can be set in motion by about 1,000 quanta of energy and 1,000 molecules of matter. This is an exciting thing to contemplate, and if correct deserves repetition and study.

EVALUATION OF PREVIOUS MEASUREMENTS

Since Langley's day eleven investigators have tried to redetermine the minimum energy necessary for vision. Not all of these workers made energy measurements; some relied on computations, some measured star magnitudes, and a few actually measured the energy required. The reason is that energy measurements are simple to describe but not easy to make.

As one studies these papers, one finds that most of them contain obvious errors which render the measurements unreliable. I have gone into this in some detail [13], and the interested reader can examine the analysis himself. Suffice it to say that three investigations [2] [4] [16] are free from what can now be recognized as serious error. The results found by these investigators are given in Table I.

TABLE I

Minimum Energy for Vision

Wavelength	*Energy*	*No. of quanta*	*Source*
$m\mu$	ergs		
505	$0.66\text{–}1.17 \times 10^{-10}$	17–30*	Chariton and Lea [4]
507	$1.3\text{–}2.6 \times 10^{-10}$	34–68	von Kries and Eyster [16]
530	$1.5\text{–}3.3 \times 10^{-10}$	40–90	Barnes and Czerny [2]

* For inexperienced observers.

From Table I it is apparent that all three agree in giving the energy a smaller value than did Langley. The energies are now in ten-billionths instead of in billionths of an erg. Even so, the measurements differ among themselves by a factor of about 5. This is too large a difference to ignore and requires investigation. Moreover, in the last twenty years the physiological and physical conditions required for maximum visual efficiency have become clearly defined [8] so that it is now possible to say precisely what the best circumstances are for determining the minimum energy.

These reasons for repeating the measurements were apparent at least eight years ago, and we resolved to make them. However, the obvious difficulties of absolute energy measurements caused us to procrastinate, and it was only after our work on the vision of the nocturnal owl was completed [12] that it became imperative to perform the threshold measurements. These were undertaken and completed by Dr. Simon Shlaer, Dr. Maurice H. Pirenne, and myself [13].

OPTIMUM PHYSIOLOGICAL CONDITIONS

To make the measurements properly one must use the most favorable physiological and physical conditions. The physiological conditions require the control of five factors. The first is dark adaptation. The eye increases in sensitivity as it stays in the dark [22] [10]. Following exposure to ordinary artificial illumination, a stay in the dark may easily result in an increase in our visual sensitivity by a factor of 1,000. But after exposures to sunlight, dark adaptation may bring about an increase in sensitivity as high as a millionfold. Dark adaptation is familiar to those who have tried to build amateur dark rooms; what looked like a light-tight room at the beginning becomes full of light leaks as one stays in the room. It is the same phenomenon that shows up as we enter a darkened theater from open daylight. At first we hardly see enough to follow the usher; but after half an hour of sitting inside, we find that the brightness of the interior has become startlingly greater. This change is in our retinas and requires at least half an hour for completion. To achieve maximal sensitivity one must allow at least three-fourths of an hour or even an hour in the dark.

After complete dark adaptation our vision is different from everyday experience. This is because we have two kinds of vision, day vision and night vision [24]. With day vision we see color, fine shades and shadows, and good detail. With night vision our detail and brightness discriminations are poor, and we make no color distinctions; but our sensitivity to light is tremendously greater than with day vision.

A pertinent aspect of night vision is its location. At ordinary and high illuminations the best part of the eye is in its center, and to see something clearly we look at it directly. With night vision the center of the eye is almost blind and our maximal visual sensitivity is off-center. This has been known for years by astronomers, and a hundred years ago Arago (1) instructed his colleagues that if they wished to see a dim star, they should not look at it but at a near-by star. The most sensitive region for night vision is about 20° off-center, and we used this locality in our measurements.

The third factor that enters into the measurements is the area of the test object. A large object may be seen even if it is dimly illuminated, whereas a small object can be seen only if it is brightly illuminated. Between area and light intensity there is an approximately reciprocal relationship (25) which has a minimum, so that an area subtending a visual angle of about 10 minutes in diameter is the most efficient stimulus. The moon is nearly 30 minutes in diameter and the circular area which we used and viewed 20° off-center has $1/3$ the diameter of the moon's disc.

A fourth condition depends on the relationship between exposure and the light intensity necessary to produce a visual stimulus. In photography it is common knowledge that on a dim day one must give a long exposure, whereas on a bright day a short exposure is all that is necessary. For producing a visual effect there is a similar reciprocal relation between time and intensity, which, however, is strictly true only for exposures below $1/60$ of a second. Therefore, to be quite certain of falling within this region of maximum efficiency, we used in our experiments an exposure of $1/1000$ of a second.

The final physiological condition involves the color of the light. The eye is not equally sensitive to all parts of the spectrum, and in this distribution of sensitivity the night eye differs from the day eye. The region of maximum sensitivity of the day eye is in the yellow-green at 550 mμ, while the region of maximum sensitivity of the completely dark-adapted night eye is at 510 mμ. This color is blue-green at high intensities; at the intensites used in night vision it appears colorless because our night vision is not capable of discrimina

To summarize: in order to use light most effectively, the eye must be thoroughly dark adapted, it should view a small field 10 minutes in diameter placed 20° off-center with an exposure of 1/1000 of a second, and the light should have a wave length of 510 mμ.

APPARATUS AND MANIPULATIONS

To secure these arrangements we devised the apparatus shown in Figure 34. This is a top view and diagrammatic

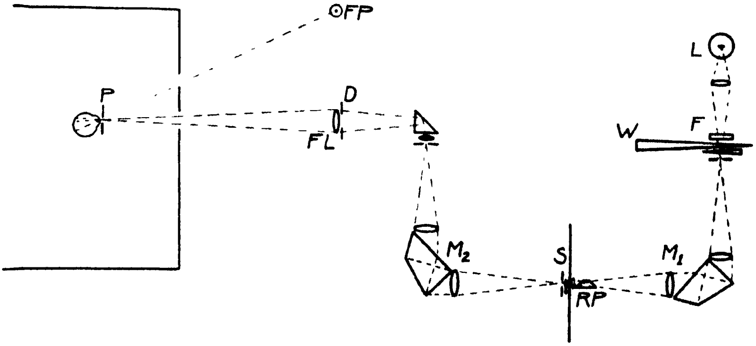


FIG. 34. Optical system for measuring minimum energy necessary for vision. The eye at the pupil *P* fixates the red point *FP* and observes the test field formed by the lens *FL* and the diaphragm *D*. The light for this field comes from the lamp *L* through the neutral filter *F* and wedge *W*, through the double monochromator M_1M_2 and is controlled by the shutter *S*.

only; its details have been described elsewhere [13]. Light from a ribbon filament lamp *L*, is focused by a lens on the slit of a spectroscope (monochromator M_1), and forms a spectrum in the plane of the slit marked *S*. This slit is set in such a way that only light of wave length 510 mμ passes through it and goes through another spectroscope, M_2, which purifies it so that there is nothing else but 510 mμ leaving the system. The light is then reflected by the prism and is focused by means of the lens, *FL*, on the cornea of the eye as it looks through the artificial pupil *P*. What the eye sees is the lens *FL* completely illuminated by monochromatic light, and the diaphragm *D* limits the size of this stimulus to a 10-minute circle.

The image of this small disc of light must now be brought to a spot on the retina 20° off-center, and arrangements must be made to have it returned to the identical spot for each measurement. This means fixing the eye in space, and is accomplished in two steps. A hard impression is made of the subject's upper jaw, and is put into a fixed position in the apparatus. By biting on this impression the head is stabilized. However, this still leaves the eye free to roam. It is kept stationary by having it look at a tiny red fixation point, *FP*, which is 20° away from the center of the diaphragm *D*. Thus the wave length, the size, and position of the stimulus are controlled.

The intensity of the light and its exposure are controlled in the following way. *F* is a series of neutral glass filters which cut down the illumination in large steps, say by 1/10 or 1/100. For continuous alterations of the intensity we use the neutral glass wedge, *W*, which by being moved in and out decreases or increases the light that passes through it. The exposure of 1/1000 of a second is achieved by the shutter *S*, the details of which need not be gone into here. It is so actuated that the observer who sits in a dark cubicle within the dark room merely presses a button and the shutter gives an exposure of 1/1000 of a second.

In the place of the eye, there may be put a thermopile which, when connected to a galvanometer and properly calibrated, records the energy incident upon it.

Two people are required for the measurements, an operator and an observer. The operator controls the light intensities, keeps the current constant, and gets everything ready for the observer. The observer merely sits comfortably in a curtained cubicle in the dark room with his eye near the exit pupil of the apparatus. When he is informed that all is ready, he looks directly at the little red light, and when he deems the moment just right, he presses the button which permits a flash of light to pass through the apparatus. He then merely records whether or not he has seen a flash of light with his peripheral vision. He has no previous knowledge of whether the light will be bright or dim or whether there will be any light there at all. He merely reports yes or no by ringing a bell, since his mouth is otherwise occupied.

After each flash the operator changes the intensity, and when he is ready he tells the observer to go on. The observer again releases a flash and records his observation. This is continued until the operator is satisfied that he knows what the observer's threshold is, or until he has the requisite number of observations from which to determine it. In general we consider that an observer's threshold has been reached when he sees a given intensity of light 60 per cent of the time.

NEW MEASUREMENTS

It is apparent that training and skill are required to be an observer; we used only seven subjects, all of whom were experienced in this kind of experimentation. Over a period of a year and a half we made measurements with them several times, with the results shown in Table II.

TABLE II

Minimum Energy for Vision

Each datum is the result of many measurements during a single experimental period, and is the energy which can be seen with 60 per cent frequency.
$\lambda = 510\ m\mu$; $h\nu = 3.84 \times 10^{-12}$ erg.

Observer	*Energy* ergs $\times 10^{10}$	*No. of quanta*	*Observer*	*Energy* ergs $\times 10^{10}$	*No. of quanta*
S. H.	4.83	126	C. D. H.	2.50	65
	5.18	135		2.92	76
	4.11	107		2.23	58
	3.34	87		2.23	58
	3.03	79	M. S.	3.31	81
	4.72	123		4.30	112
	5.68	148	S. R. F.	4.61	120
S. S.	3.03	79	A. F. B.	3.19	83
	2.07	54	M. H. P.	3.03	79
	2.15	56		3.19	83
	2.38	62		5.30	138
	3.69	96			
	3.80	99			
	3.99	104			

Two things come from an examination of this table. First, it is encouraging that the order of magnitude of our measurements is the same as that of the best three previous determinations already shown in Table I. All our observations lie between 2.1 and 5.7 × 10^{-10} erg. They are about twice as large as the previous measurements, and there is a certain amount of overlap between them and the others. Nevertheless, one can confidently consider them as confirming one another.

Since the energy of a single quantum of blue-green light is 3.84 × 10^{-12} erg, our measurements represent thresholds between 54 and 148 quanta. The second thing apparent from Table II is that an individual's threshold varies over a period of time. Mine (S. H.) varies between 79 and 148 quanta; Shlaer's (S. S.) between 54 and 104; and Pirenne's (M. H. P.) between 79 and 138. In other words a person's threshold may vary by a factor of 2, and this agrees with a great deal of previous work on thresholds.

The original question raised at the beginning of our inquiry has now been answered. In order for us to see a light, we must receive from it between 54 and 148 quanta. However, this answer far from exhausts the interest of the measurements. In fact, their real significance has not been even hinted at, and it is to this that we may now turn.

QUANTUM LOSSES

The 54 to 148 quanta necessary for seeing represent the energy incident on the cornea, and we have tacitly made the assumption that they represent the actual energy required for initiating a visual act in the retina. This is not correct. There are at least three corrections that must be made before these numbers acquire meaning in terms of chemical processes in the retina.

Figure 35 presents a horizontal section of the human eye with its main landmarks indicated. The drawing shows that light must pass through several structures before it reaches the retina. The first solid surface which the light reaches is the smooth and wet cornea. From such a surface it suffers a reflection of about 4 per cent. This means that of the minimum 54 quanta 2 are thrown away before entering the eye.

However, after entering the eye the light passes several barriers. As it leaves the cornea it enters the aqueous humor; this is a fluid, undoubtedly less dense than the cornea, and as a result the light will suffer another reflection here. After this it reaches the front surface of the very dense lens, and then the rear surface, to enter and pass through an inch of jelly called the vitreous humor before it impinges on the retina. At all these surfaces it will suffer reflection, and it will also be absorbed and scattered by the ocular media which contain protein.

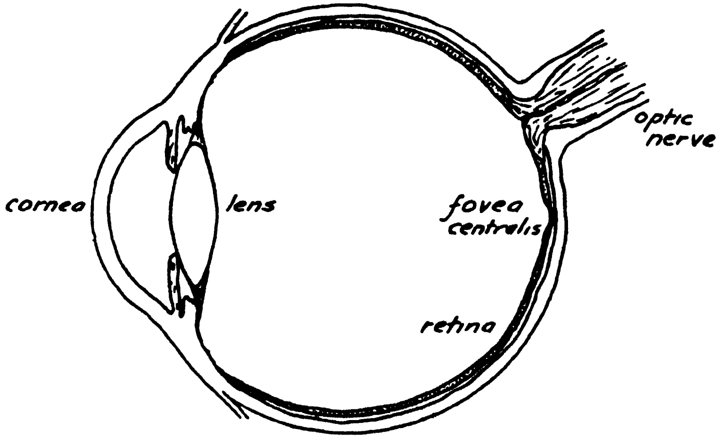


FIG. 35. Diagram of a horizontal section through the human eye. It is evident that light in going successively through the cornea, the aqueous humor in the anterior chamber, the lens, and the vitreous humor in the posterior chamber, must suffer some reflections and absorptions before it reaches the retina where it does its photochemical work.

It used to be thought that the eye media are completely transparent. This must be partly true, because we really do see through it with excellent precision. But it is not completely transparant, as Roggenbau and Wetthauer [23] have shown with cattle eyes. One may then inquire what fraction of the light which enters the cornea finally reaches the retina. This has been measured by Ludvigh and McCarthy [19] with human eyes, and their averaged measurements for the 21-year-old eye are shown in Figure 36.

From Figure 36 it is evident that for violet light only about 10 per cent of the incident light reaches the retina, while for red the fraction is about 70 per cent. For the particular light in which we are interested, namely 510 mμ, almost exactly 50 per cent of the entering light reaches the retina. In an older eye the fraction is even less.

From all this it is clear that of the 54 quanta which fall on the cornea only 26 finally reach the retina. It seems almost profane to be so free with these few precious quanta, yet this is not the end.

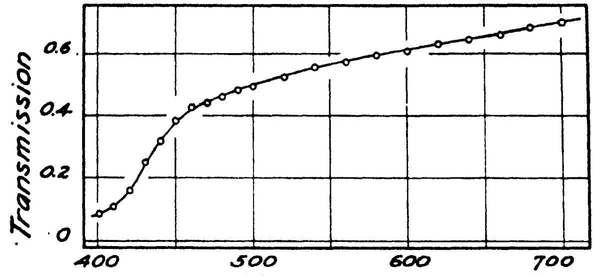


FIG. 36. Transmission of light through the human eye. The measurements are by Ludvigh and McCarthy [19] and give the fraction of incident light which finally reaches the retina. This fraction varies with the wave length of the light.

Figure 37 shows an enlarged diagram of the retina. The light comes from the left and passes through the nerve cells to fall on the layer of sensitive elements to the right. These are the rods and cones, of which three each are drawn, while the rest are indicated by horizontal lines. These retinal elements are closely packed one next to the other, with no spaces between. The cones are of no interest to us here, but the rods—shown by the dark terminal pieces—are the elements concerned in the present problem, and are present to the extent of 150,000 per square millimeter [21].

In order to be effective, the light must be absorbed by the terminal segments of the rods. Here takes place the photochemical conversion of light into sensory impulses, which pass through the bipolar cells into the ganglion cells and through the brain to the occipital cortex. Any light which is not absorbed by the rods passes through the retina and is absorbed by the densely black pigment cells at the extreme right of Figure 37.

In our measurements we have found that when 26 quanta fall on the retina they constitute a minimal visual stimulus. Clearly all of this light cannot be absorbed by the sensitive elements, because they would have to be practically black, and we know they are not. The question then is what fraction of this light is absorbed by the rods for chemical work, and what fraction goes through unabsorbed.

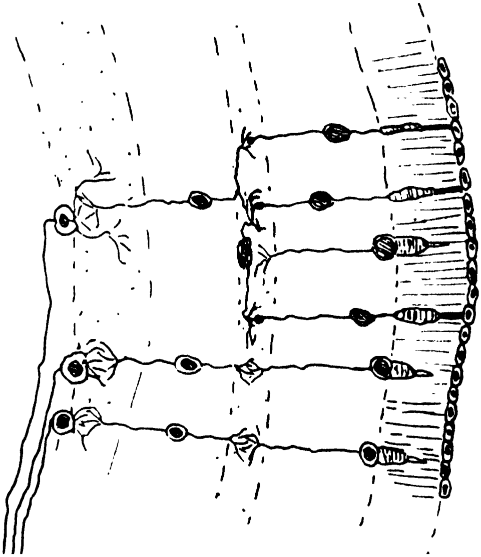


Fig. 37. Diagram of the structure of the retina. Light comes from the left, passes through the nerve cells, and impinges on the rods and cones, which are the sensitive elements. The cones are for day vision, and are shown as the stubby elements, while the rods are for night vision, and are drawn as the thinner and longer elements with well-defined terminal segments touching the layer of very black pigment cells at the extreme right. These terminal segments contain visual purple and in them occur the initial photochemical events involved in seeing light at low illuminations.

It is the visual purple in the terminal segments of the rods which absorbs the light, and one can determine the amount it absorbs by ascertaining how much visual purple a human retina contains. Such measurements were made by Koenig [15], who computed that if this quantity were spread over the whole retina, it would absorb only about 5 per cent of light of 510 mμ. This is a small value indeed, and would seem almost unacceptable if it were not that Wald [27] has found similar quantities in the retinas of the rabbit and the rat, namely, 4 per cent and 13 per cent respectively.

Even so, these figures are probably too low. First, it is not likely that all the visual purple in an eye is extracted by the usual procedures. And second, the figures rest on the supposition that visual purple is evenly distributed over the retina. This is certainly not true, because it is missing in the fovea; and in the regions outside of the fovea its concentration undoubtedly increases to a maximum in the region about 20° off-center corresponding to the high density of rod population in this region. The value of 5 per cent, however, may be considered as setting the lower limit for retinal absorption.

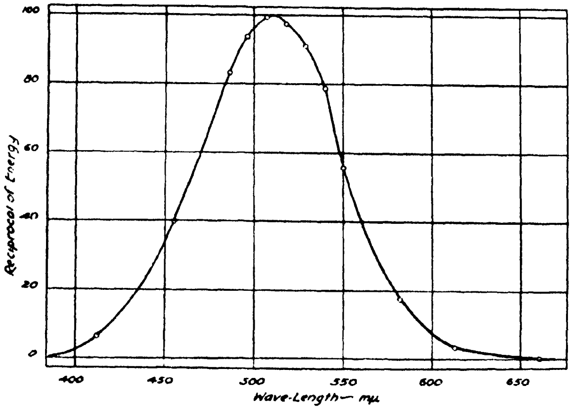


FIG. 38. Sensibility distribution of the eye in the spectrum near the threshold of vision. It is apparent that the eye is most sensitive at 510 mμ and least sensitive at the two ends of the spectrum.

If one accepts this value, or even 10 per cent, as representing retinal absorption, it means that of the 26 quanta falling on the retina no more than 2 or 3 are absorbed by the rods to start the events which end in our seeing light. This is so small a number of quanta that it seemed imperative to determine retinal absorption in a way wholly independent from that used by Koenig and by Wald. We were able to do this by considering the influence of concentration on the absorption spectrum of a substance and the relation which this bears to the photochemical effectiveness of the light incident on it.

SPECTRUM SENSIBILITY AND VISUAL PURPLE

I have referred previously to the sensitivity distribution of the eye in the spectrum. Figure 38 shows the data which Williams and I [14] obtained in 1922, and records the average sensitivity of 48 people to the spectrum at an illumination only slightly above the threshold. The results have since been confirmed by Weaver [28] with 14 people. From this curve one can see that the maximum sensitivity is at 510 mμ, and that the eye is relatively insensitive both to the red and to the blue parts of the spectrum.

Why does the eye possess such a sensitivity distribution? Or to put it the other way around, why is light of 510 mμ so much more effective as a stimulating agent than red and blue? Since light can be effective chemically only when it is absorbed, we must suppose that 510 mμ is more effective than the ends of the spectrum because it is absorbed to a greater extent by the light-sensitive substance in the retina. From the shape of the spectrum sensitivity curve in Figure 38 it is possible to derive the precise concentration and absorption of the sensitive material in the retina provided the absorption spectrum of the material is known.

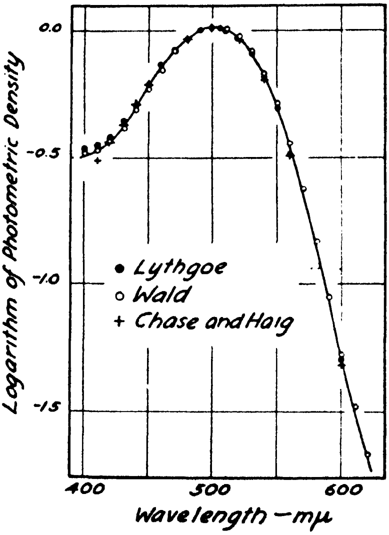


FIG. 39. Absorption spectrum of frog's visual purple. The data from the three sources have been made equal at 500 mμ.

The sensitive substance in the terminal segments of the rods is visual purple, a material about which a great deal has become known since it was discovered by Boll [3] in 1876 and extracted by Kühne [17] a year later. For example, it is known, among other things [9], that visual purple is a protein, whose molecular weight [11] is about 270,000, and that it has a carotenoid prosthetic group derived from vitamin A which is responsible for its sensitivity to light [25]. More important still, it has been possible to purify visual purple, and to measure its absorption spectrum with precision.

Figure 39 shows the absorption spectrum of frog's visual purple as determined independently three times, by Lythgoe [20] in London, by Wald [27] at Harvard, and by Chase and Haig [5] in our laboratory. The three determinations agree well, so that the absorption spectrum may be considered as an established characteristic of the pure substance.

Strictly speaking, the curve in Figure 39 represents the theoretical absorption of visual purple. It records the absorption of an infinitely thin or of an infinitely dilute solution. However, we are not concerned with either of these two theoretical conditions but with real concentrations and thicknesses. From the data of Figure 39 it is possible to prepare a series of percentage absorption spectra for different concentrations of visual purple by simple computation in terms of well-known laws relating concentration and absorption. Such a series of computed curves is shown in Figure 40. For comparison among themselves, the maxima of the curves have been made equal to unity, but their actual values are shown in the diagram. The important aspect of Figure 40 is that the width of the percentage absorption band increases as the concentration of visual purple increases. Carried to its logical conclusion, this means that if one were to get a thick enough layer of visual purple, its absorption spectrum would be so wide that it would spread over the entire visible spectrum, and would thus be opaque visually. Those who have worked with visual purple know that it will be years before so much pure substance can be prepared.

Applied to the spectral sensitivity of the rods, Figure 40 tells us not only that the shape of the curve in Figure 38 depends on the absorption by visual purple of the different parts of the spectrum, but that its width must correspond to some specific concentration of visual purple in the rods. If the concentration is high, the spectrum sensibility curve will be wide; if the concentration is low, the sensibility curve will be narrow. By comparing the spectrum sensibility curve of Figure 38 with the various absorption spectra in Figure 40, it should be possible to decide on the precise concentration of visual purple actually present in the retinal rods.

Before the spectrum sensibility data of Figure 38 can be used for this comparison, they must undergo two transformations. The first is to change their location from the cornea to the retina, and this involves correction for ocular media transmission by means of the data of Ludvigh and McCarthy. The second is to convert them from an energy basis to a quantum basis as suggested by Dartnall and Goodeve [6]. Spectrum sensibility records the photochemical effectiveness, and it is the number of quanta absorbed which determines the photochemical effect and not just the energy absorbed. This conversion is achieved by dividing the spectrum sensibility data by the corresponding wave lengths.

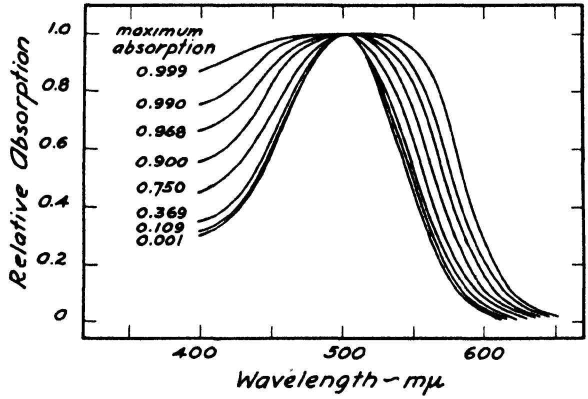


FIG. 40. Percentage absorption spectra of various concentrations of visual purple. For convenience in comparing the shapes of the curves, their maxima have all been equated to 1 and superimposed. The actual fraction absorbed at the maximum is shown for each curve. It is apparent that with increasing concentration the absorption curve steadily increases in width.

Transformed in this way our old data are shown as circles in Figure 41. Included also in Figure 41 are two absorption curves derived as were those in Figure 40. The wider of the two curves represents a concentration of visual purple such that its maximum absorption is 20 per cent, while the narrower curve is a concentration of visual purple whose maximum absorption is 5 per cent. The 5 per cent maximal absorption curve describes the points fairly well. The 20 per cent curve though it touches many of the points is just a little too wide. The curve which best describes the data is a 10 per cent maximal absorption one; this is not shown in the figure, to avoid confusion.

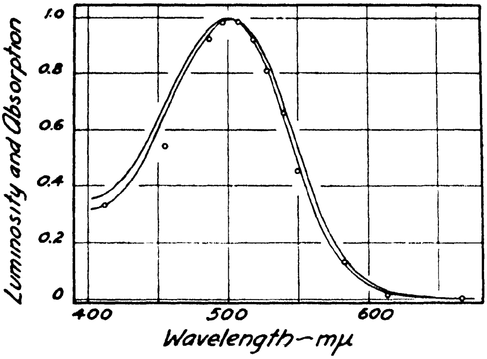


FIG. 41. Comparison of scotopic luminosity at the retina with visual purple absorption. The points are the data of Hecht and Williams [14] corrected for quantum effectiveness and ocular media transmission. The curves are the percentage absorption spectra of visual purple; the upper curve represents 20 per cent maximal absorption, and the lower one 5 per cent maximal absorption. All curves have been made equal to 1 at the maximum, 500 mμ, for ease in comparison.

The significant point of these comparisons is that the 5 per cent and 10 per cent absorptions are of the same order of magnitude as those found from direct measurements by Koenig and by Wald. We thus have a double check of this dubious item in the correction of our quantum measurements. In order to be entirely safe, we may take 20 per cent as the upper limit for the absorption of light of 510 mμ by the visual purple in the eye at the threshold during complete dark adaptation.

QUANTA IN THE RODS

We may now return to the original problem of how many of the 54 to 148 quanta incident at the cornea are ultimately used in the retina for the chemical work needed in producing a just perceptible visual effect. Four per cent is immediately reflected by the first surface; very nearly 50 per cent is absorbed by the lens and by the ocular media; and of the energy which impinges on the retinal layer, a maximum of 20 per cent is absorbed by the visual purple itself. These corrections when applied to a range of 54 to 148 quanta convert them into 5 to 14 quanta absorbed by the retinal cells.

These are small numbers. However, they will be decreased still more when we consider one additional factor in the measurements.

It will be recalled that the image formed on the retina by the flash of light occupies a circle whose diameter is 10 minutes of arc. Since this region of the retina contains about 150,000 rods per square millimeter, computation shows that our 10-minute disc covers about 500 rods. Only 5 to 14 quanta are absorbed by these 500 rods; it is therefore very unlikely that any one rod will take up more than one quantum. Simple statistical considerations show that if 7 quanta are absorbed by 500 rods, there is only a 4 per cent probability that a single rod will take up more than one quantum. To put it the other way, the chances are 96 out of 100 that each quantum is separately absorbed by one rod, and that each of 5 to 14 retinal rod cells is activated by one quantum.

In terms of the photochemical equivalence law, each quantum changes one molecule of visual purple. The data then say that for us to see a just perceptible flash of light, there need be transformed photochemically only one molecule of visual purple in each of 5 to 14 rod cells.

FLUCTUATION OF DISCRETE, RANDOM, AND INDEPENDENT EVENTS

We have now answered for the second time the original question with which this study began. And our answer is almost embarrassing. Five to 14 quanta or molecules constitute a number of physical and chemical events so small that one is hesitant to accept it. For this reason we tried to find an independent way in which to test the answer—a way totally different from what had been done up to now. Such a method is available, and its essential point is that it depends for its effectiveness on the very fact that the number of events is small.

In our measurements when we say that the energy per flash is, say, 10 quanta at the retina, we derive this value as the result first of a calibration, and then of a computation. The calibration is made with a thermopile and a galvanometer, and records an energy density. It takes several seconds to make the measurement, and the answer is, let us say, 10,000 quanta per second. Then, knowing that our shutter gives a flash of $\frac{1}{1000}$ of a second, we compute the flash to deliver 10 quanta. It is thus a computed average value.

In reality, however, each flash will not always deliver this average number. Because quantum emissions and absorptions are discrete, random, and independent events their occurrence will show a certain kind of fluctuation. Sometimes a flash will yield more than the average number of quanta and sometimes fewer. And the way in which the number of quanta per flash will occur in a series of flashes will follow a probability distribution developed by Poisson and known by his name [7].

Assume that in a series of flashes the average number of quanta per flash is a. Then the Poisson distribution will tell us the frequency with which any number of quanta n will appear in the series of flashes. If P_n is the probability that a flash will yield n quanta, its value is given by the simple equation

$$P_n = \frac{a^n e^{-a}}{n!}$$

in which e is the base of natural logarithms. Values of P_n for various values of a and n are available in printed tables [7].

This series may be applied to vision in the following way: Suppose that it takes n quanta to activate the retina for us to see a flash of light. Then in any series of flashes which averages a quanta per flash, we will see the flash whenever it delivers n quanta even when n is greater than a. Moreover, we shall see those flashes which yield more than n quanta.

For example, if a series of exposures averages 10 quanta per flash, and if we need 12 quanta to activate the retina then every time an exposure happens to yield 12 or more quanta, we will see it as a flash of light.

From the published Poisson distributions it is simple to compute the probability that *n* or more quanta will be delivered for any average value of *a*. The results of such a computation are shown in Figure 42. From the curves in Figure 42 one can read off the probability of getting *n* or more quanta for a series of average values. For example, if the average number is 1 (on the graph in Figure 42, log 1 = 0) then there is over 60 per cent chance that any flash will deliver one or more quanta. But there is also about 25 per cent chance that it will deliver two or more quanta, and there is almost a 10 per cent chance that it will deliver 3 or more quanta. The thing to note in Figure 42 is that for each value of *n* the shape of the curve is different from every other value when the number *n* is small. The curve which relates the frequency of getting *n* or more with the average number becomes steeper as the value of *n* increases. The curves for the values of *n* differ so much from each other that if the probability distribution can be determined by experiment, its shape will immediately reveal the value of *n* which corresponds to it.

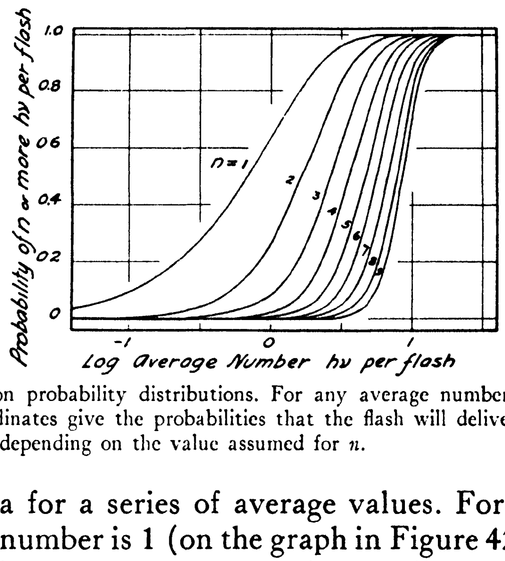


FIG. 42. Poisson probability distributions. For any average number of quanta (hν) per flash, the ordinates give the probabilities that the flash will deliver to the retina *n* or more quanta, depending on the value assumed for *n*.

FREQUENCY OF SEEING

Experimentally this means that when a flash of light is given repeatedly to the eye, there will be a certain frequency with which the flash will be seen as a light. If the average energy is high, the flash will be seen nearly every time. If the energy is low, the flash will not be seen every time but only that percentage of the time when it delivers the requisite number of quanta. All that we need is to record the frequency of seeing a series of flashes of different intensities. The results, when plotted as frequency against the logarithm of the relative intensities, will resemble one of the several curves in Figure 42, and the particular curve which fits it best will yield the value of n corresponding to the number of quanta needed for recognizing a flash of light. To make the experiments in this way, one needs no monochromators, no thermopiles, and no galvanometers; the only measurement required is the relative intensities of the light flashes.

Actually our experiments were made with the same apparatus as before, but this was purely for convenience. The intensities were varied in a deliberately random fashion but each intensity was presented 50 times in the course of an experiment, and for each flash the observer merely recorded whether he saw it or not. Experiments like these are fairly rigorous and only skilled observers are useful. Comfort is of great importance. The subject must be relaxed and fixation need not be continuous. Moreover, the observer must watch for any sign of fatigue and call the experiment off if it persists. The observer controls the occurrence of the flash by pressing a button when he is ready. All the rest of the work is done by the experimenter who varies the intensities and makes the recordings.

The first results secured for the three of us are shown in Figure 43. In the figure the abscissas give the logarithm of the average number of quanta per flash which fall on the cornea only because we happen to have measured them. The real values are unnecessary; only the relative values are required. Through the points in Figure 43 there have been drawn Poisson distribution curves from Figure 42. The measurements are best fitted by such distributions when *n* is 5, 6, and 7 quanta per flash. In subsequent experiments the value of *n* has gone as high as 8. No special methods are needed to determine which curve best fits the data, because smaller and larger values of *n* are easily excluded by the simplest visual comparison.

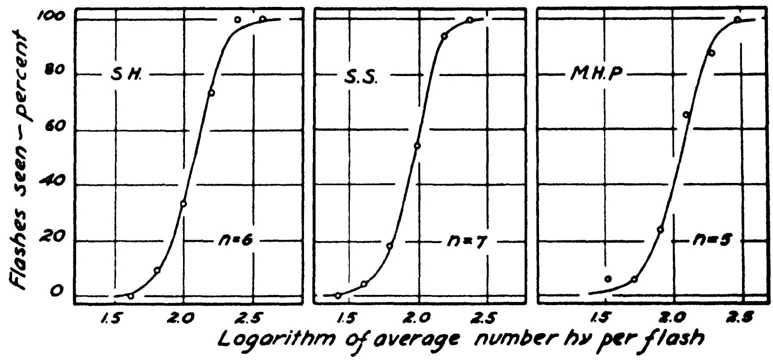


FIG. 43. Relation between the average energy content of a flash of light (in number of hν) and the frequency with which it is seen by three observers. Each point represents 50 flashes, except for S. H. where the number is 35. The curves are the Poisson distribution of Fig. 42 for *n* values of 5, 6, 7.

Here, therefore, by a complete independent method we arrive at the same answer to our original question as we secured by straightforward physical measurements. One can hardly doubt that the necessary energy for a minimum visual effect lies between 5 and 8 quanta.*

TWO QUESTIONS

Before concluding, it may be of interest to discuss two points which undoubtedly occur to the inquiring reader.

---

* In a recent summary of "Optical Physics in the U. S. S. R.," by S. Frisch (*Jour. Opt. Soc. Amer.*, *33*, 637, 1943), it is reported that Vavilov has just determined the threshold of vision by direct physical means, and simultaneously by frequency of seeing measurements. Frisch states that the direct measurements give the threshold as 200 quanta at the cornea, while the fluctuation method puts it as 25 quanta per second at the retina (the "per second" is probably superfluous). The direct value is somewhat large, but is near enough our results and those of von Kries and Eyster and Barnes and Czerny to be considered as reasonable confirmation. The retinal values, however, are definitely too large by a factor of about 3. We have not been able to see these recent papers by Vavilov. His earlier paper with Brumberg (*Bull. Acad. Sci., U. S. S. R.*, 1933, 919) has already been referred to [13] and evaluated. A fresh use of fluctuations has just been reported by Pirenne, in a study of the binocular and uniocular visual threshold (*Nature*, *152*, 698, 1943).

The first is really a question. Why do we need 5 quanta for a stimulus and why is one not enough? No certain answer can be given to this, but a possible one suggests itself.

The quantum effects start nerve impulses which ultimately go to the occipital cortex. Because of changes in blood pressure and in muscle tension of the eye, and because of fluctuations in the metabolism of the eye, the optic nerve, and the brain, it is conceivable that even without light a single nerve impulse might start any place along the path from retinal rod cell to the terminus in the brain. The chance of such a fortuitous single nerve impulse occurring is probably high, and the brain must be on guard to avoid confusing it with a flash of light. The chance of two such accidental impulses originating simultaneously is not nearly so high, but their occurrence is not improbable. However, the chance that five spontaneous impulses will arise simultaneously in closely related nerve fibers is completely negligible. Therefore, when five near-by fibers are simultaneously active, it can only be as the result of a light shining on the corresponding retinal locality.

The second point concerns the necessity for working with such small amounts of energy. Surely there is plenty of light available most of the time, and it seems hard to understand why the human eye has had to become so sensitive as to deal with single quanta.

The answer to this is probably purely biological. Our night vision must have developed in response to the need for seeing other animals before they could see us for their ulterior purposes. By the same token, *their* night vision developed for similar reasons. It is therefore a tribute to the power of natural selection that because of it animals became more and more sensitive until they reached the limit set up by the physical structure of light. One quantum is the smallest unit of energy, and when a retinal rod cell is capable of being brought into action by one quantum of light, the end has been reached and nothing more can be done in this direction.

IV

CHEMICAL TRANSMISSION OF NERVE IMPULSES

By OTTO LOEWI

New York University

INTRODUCTION

It is the function of the central nervous system to control the physiological activity of the body according to its needs. For the most part the central nervous system is informed about these needs by messages reaching it through the afferent nerves. It responds by sending out impulses, which eventually reach the effector organs by way of the efferent nerves and there induce the appropriate functional changes. Until recently the mechanism by which the nerve impulses are transmitted to the effector organ was not understood. All that was known was that if a nerve is stimulated, either naturally or artificially, an excitation is produced in the connected muscle tissue or other effector. This excitation is propagated within the nerve and is accompanied by an electrical wave, the action current. It was generally believed that this electrical wave traveled to the effector organ and there elicited the specific reaction. This theory, however, encountered two difficulties. In the first place, there is no tissue continuity between the nerve and the effector organ which would permit the electrical wave to pass from one to the other; they are merely contiguous, that is, the end of a nerve fiber makes only a synaptic contact with the effector organ. Secondly—and this was the greater difficulty—how is it possible for electrical waves propagated over certain fibers to increase the activity of the effector organ, whereas impulses from other so-called inhibitory nerves decrease or even stop effector activity? The theory of electrical transmission was obviously not the answer to the problem, and another solution had to be found.

Elliott in 1905 demonstrated [1] that the effects of adrenaline in the organism were identical with those of stimulation of sympathetic nerves. This led him to suggest [3] that adrenaline liberated from the endings of the sympathetic nerves during stimulation might be responsible for the effects observed. Later Dixon [2] considered the possibility of another substance being liberated by stimulation of the vagus nerve. Indeed he reported finding such a substance, but he never carried his investigation further. The reports of both Elliott and Dixon were published so inconspicuously that they escaped general attention, as well as my own. The earlier results were only rediscovered, so to speak, some years after my discovery of the chemical transmission of the nerve impulse, which was made in 1921 [4]. It was the result of a simple experiment, which I shall describe in some detail, since it is the basis of the whole theory of chemical transmission of nervous impulse.

ESTABLISHMENT OF THE THEORY OF CHEMICAL TRANSMISSION

The problem was to learn whether stimulation of a nerve liberated a chemical substance from the endings, which, in its turn, was able to produce the effect observable in nerve stimulation. The hearts of two frogs were isolated: one with the vagus and accelerator nerves connected to it, one without connecting nerves. The first heart may be designated as the donor heart, the second as the recipient. In each heart a small glass cannula was inserted containing Ringer solution. For a short period the vagus nerve connected with the donor heart was stimulated with the usual effect: decrease of frequency and amplitude of the heart beat. Then the liquid content of this donor heart was transferred to the recipient heart, with the result that the latter behaved exactly as the former had during the vagus stimulation. When, instead of the vagus, the accelerator nerve of the donor heart was stimulated, and the content later transferred to the recipient heart, the effect was again identical: in this instance increase of frequency and amplitude of the heart beat. Here was the proof that the nerves do not influence the heart directly, but that chemical substances liberated from their endings by stimulation cause the characteristic modification of the heart action. In other words, a nervous impulse is transmitted to the cardiac tissue of the heart indirectly by chemical substances which we may call chemical transmitters.

It seemed probable that the mechanism of chemical transmission was not limited to the special case of the heart but would be found in other effector organs as well. We shall soon see that this assumption was correct; but first, in view of what is to follow, let us consider the nature of the substances released by the nerve terminals following stimulation.

I should like to begin by stating how I succeeded in identifying the transmitter of vagus impulses, which I then called *Vagusstoff*. At that time only one substance in the organism, choline, was known to produce an effect similar—at least in certain respects—to that of vagus stimulation. The Vagusstoff could not be choline, since the effect of choline introduced into the heart lasts a long time, whereas the effect of the Vagusstoff fades quickly. On the basis of Reid Hunt's discovery [5] that some cholinesters are much more effective than choline itself, and knowing that their action is transitory, I had already assumed by 1922 that the Vagusstoff might be a cholinester. This assumption later proved to be true, for it was possible to show that the course and mode of action of the Vagusstoff and of a certain cholinester, acetylcholine, are indistinguishable [6]. The rapid disappearance of activity characteristic of both of these substances proved to be caused by a hydrolyzing enzyme present in the heart, later named cholinesterase. This enzyme splits and renders both the Vagusstoff and the acetylcholine ineffective in the same degree.

Another indication that the two substances were identical was the following. It had long been known that the transitory effect of vagus stimulation is greatly prolonged by the previous application of very small doses of the alkaloid, eserine. The discovery of the action of cholinesterase suggested an inquiry as to whether this prolongation might not be due to an inhibitory effect of eserine on cholinesterase. This proved to be, in fact, the case. After an application of eserine the action of the Vagusstoff and acetylcholine, like the effect of vagus stimulation, is indefinitely prolonged. The discovery of the esterase-inhibiting action of eserine made possible the later extension of the proof of the occurrence of neurochemical transmission to a very wide range of effects.

This work and the fact that two excellent British scientists, Dale and Dudley [8], later succeeded in isolating the same substance from animal organs, leaves no doubt that the Vagusstoff is identical with acetylcholine.

The substance released by stimulation of the cardioaccelerator nerves was identified in 1936 as adrenaline [9]. In this country it still is frequently called sympathine, by a misunderstanding. Cannon [10] for various reasons has assumed that the sympathetic transmitter, after having been released from the nerve endings, combines in the organ where it was released with special substances. He agrees, however, that the substance as released from the sympathetic nerve endings (with which alone we are concerned) is, in fact, adrenaline.

GENERAL APPLICABILITY OF THE THEORY OF CHEMICAL TRANSMISSION

With this information we may next consider the general applicability of the neurochemical mechanism of transmission, as discovered in the frog's heart. In the first place it was possible to show that this mechanism also functions in the transmission of the nervous impulse in mammalian hearts. The nerves controlling the heart action, the vagus and the accelerator, belong to two divisions of the autonomic nervous system: the vagus to the parasympathetic, the accelerator to the sympathetic. Hence it was to be expected that the neurochemical mechanism of transmission would apply to the other nerves of the autonomic system as well. This theory was tested for a wide range of cases. A typical instance is found in the transmission of the impulse from the oculomotor nerve of the eye to the circular sphincter muscle of the iris [11]. On the assumption that if acetylcholine were to be liberated from the endings of the oculomotor nerve some might diffuse into the aqueous humor, Engelhart in my laboratory proceeded to eserinize both eyes of a rabbit in order to prevent the splitting of acetylcholine incidentally liberated. One eye was exposed for a time to strong light, the other kept in the dark. Then the aqueous humor of each eye was withdrawn and tested on the frog's heart. The result was that the aqueous humor obtained from the eye exposed to light acted on the heart exactly like acetylcholine, whereas that of the other eye was without effect. Hence it was definitely proved that the reflex constriction of the muscle fibers in the sphincter of the iris produced by the light stimulus is due to the release of acetylcholine from the oculomotor nerve endings. The neurochemical mechanism obviously holds for all constrictions of the sphincter muscle in the iris.

In brief, the mode of transmission of impulses from the postganglionic autonomic nerves to their respective effector organs has been proved, in all cases studied, to be through the release of a chemical transmitter. As to the nature of the chemical transmitters, most of the postganglionic parasympathetic nerves release acetylcholine, most of the sympathetic ones release adrenaline. In a few instances, however, the reverse has been found to be true. Following a suggestion of Dale [12] we therefore distinguish "cholinergic" and "adrenergic" fibers, irrespective of their anatomical origin.

So far in this discussion, only the mechanism of transmission from the postganglionic autonomic nerves to their respective effector organs has been considered. We shall now see that neurochemical transmission occurs far more widely.

CHEMICAL TRANSMISSION IN SLOW AND RAPID ACTIVITY

In the studies described above the effect of stimulation of the nerves on their effector organs is relatively slow; it starts slowly, gradually reaches a maximum, and then slowly fades. It seemed theoretically probable that in these cases a chemical process might take place, consisting in the release and disappearance of a chemical substance. There exist, however other instances in which the effect starts and stops almost simultaneously with the stimulation as, for example, in the case of transmission from the preganglionic fibers to the ganglia or from the somatic nerves to the skeletal muscles. The speed of the action here made it seem doubtful that the transmission was effected by the liberation of a chemical substance.

Before discussing the experiments concerned with these examples of rapid transmission, it should be mentioned that acetylcholine has two points of attack within the parasympathetic nervous system. At the postganglionic synapses the action of acetylcholine closely resembles that of the alkaloid, muscarine, and hence is generally called muscarine action. On all the autonomic ganglia acetylcholine acts just as does the alkaloid, nicotine, stimulating the ganglia when small doses are used and paralyzing them in large doses. This, accordingly, is called the nicotine action.

The first investigation was made in Dale's laboratory [13]; it concerned the character of the transmission of excitation from a preganglionic nerve fiber to the superior cervical ganglion, which was perfused with Ringer solution. It was demonstrated in the first place that acetylcholine, added in low concentrations to the perfusion fluid of the ganglion, had a stimulating effect on the ganglion, whereas in strong concentrations it had a paralyzing effect; this result was exactly like that of nicotine. Furthermore it could be shown that during preganglionic stimulation acetylcholine was present in the eserinized perfusion fluid of the ganglion, in a concentration sufficient to stimulate the ganglion. This acetylcholine, as will be seen later, originates exclusively from stimulated preganglionic nerve endings. To summarize: acetylcholine is liberated by stimulation of the preganglionic fibers and is able to stimulate the ganglion. While from these facts alone it could be assumed that acetylcholine is the transmitter of impulses from the preganglionic fibers to the ganglion, the correctness of this view has been substantiated by the results of various other experiments. For instance, it has been shown that when the preganglionic fibers were stimulated to the point of exhaustion, the reaction of the ganglion ceased. At this same moment acetylcholine disappeared from the perfusion fluid. After an interval during which no stimulation of the preganglionic fibers occurred, the reaction of the ganglion returned and, at the same time, acetylcholine reappeared in the perfusion fluid.

In view of these and other findings there seems to be abundant evidence for the belief that acetylcholine is the chemical transmitter of the nerve impulses. Only one point still needs amplification: the fact that the effect of preganglionic stimulation both starts and stops when stimulation does. Assuming that acetylcholine is the transmitter, it is obvious that its effect must immediately vanish when stimulation ceases. How could this happen? We have already seen that acetylcholine is present in the perfusion fluid only when the fluid contains eserine. It may, therefore, be assumed that the rapid disappearance of the effect of nervous stimulation may be the result of a splitting of acetylcholine by cholinesterase. As a matter of fact Nachmansohn (27) demonstrated that there is present at the synapse a concentration of cholinesterase high enough to split the acetylcholine released by preganglionic stimulation within ten milliseconds.

Another instance in which the effect of nervous excitation immediately begins and disappears, and where it has also been proved that the transmission of the nervous impulse to the effector organ is of a chemical nature, may be noted in the transmission of an excitation from the spinal nerves to the striated muscle.

What are the facts here? Dale and his coworkers [26], after indirect stimulation through nerve fibers of artificially perfused muscles, found acetylcholine in the eserinized perfusion fluid. On the other hand, the direct stimulation of the muscle tissue after nervous degeneration did not release even a trace of acetylcholine, though the muscle contracted strongly. Hence the acetylcholine which, in the first experiment, was found in the perfusion fluid was proved to be liberated from the nerve endings, and was not a product of muscular activity. These results, however, did not prove that acetylcholine was the transmitter. Only a muscle twitch produced by the action of acetylcholine, instead of by nervous stimulation, could be regarded as positive proof.

This proof was established by Dale and his coworkers [24]. To imitate the sudden simultaneous release of acetylcholine everywhere in the muscle as caused by nervous stimulation, they bled the muscle and then injected a small amount of acetylcholine in the artery close to the muscle. The effect was amazing. The injected acetylcholine produced a strong quick muscle twitch very similar to that evoked by nervous stimulation. The drug, curare, and large doses of nicotine, by their well-known effect on the motor end-plate, inhibited the effect of acetylcholine exactly as they did following nervous stimulation; and previous eserinization influenced the effect of the acetylcholine injection in the same way as on nerve stimulation: i.e., instead of a single twitch there was a series of twitches forming a tetanus. This last result also demonstrated that the action of cholinesterase (which, as we have seen, is prevented by eserine) is at least one of the mechanisms responsible for the quick disappearance of the muscle twitch caused by nervous excitation as well as by injected acetylcholine. In this case too Nachmansohn [28] has shown that the concentration of cholinesterase present in parts of the muscle containing nerve endings is sufficient to account for the rapid removal of the acetylcholine released by nervous stimulation.

CHEMICAL VERSUS ELECTRICAL TRANSMISSION AT SYNAPSES

Though, as we have seen, everything points to acetylcholine being the transmitter of nervous impulses, even in those cases where the effect is very quick, as in the preganglionic and spinal nerves, some eminent neurophysiologists, especially in this country, still believe that the transmission in these cases is electrical rather than chemical. They hold that the action current caused by excitation in the first neuron, preganglionic fiber, or spinal nerve passes through the synapses where the first neuron makes contact with the second and, eventually, stimulates the effector organ, ganglion, or muscle, respectively. The main reason for their not accepting the chemical theory in these cases is the brevity of the latent period before the action of the postganglionic elements occurs. It is true that the synaptic events at the ganglionic cells and the motor end-plates are very brief. According to Eccles [18], in the neurons of the superior cervical ganglion the latent period ranges from 2 to 4 thousandths of a second, and in the fibers of some skeletal muscles it can be even shorter, in fact about one half that. It has been claimed that this period is too short for the liberation action and disappearance of acetylcholine. This, however, is merely a guess. Almost nothing is known about the possible speed of chemical reactions within the organism and of the mechanism involved. Consider, for instance, the fact that a bird's heart may beat 450 times a minute and during each beat the metabolism swings from one extreme to the other. So long as we know so little of the factors possibly involved *in vivo* that cannot be imitated *in vitro,* I do not consider the time factor a serious argument against the chemical theory of transmission.

At this point I should like to call attention to an important analogy disclosed by some of the facts just discussed. As has been shown, acetylcholine affects the ganglionic cells and the motor end-plate similarly, stimulating them in small doses and paralyzing them in large ones. So does nicotine. Curare, although having the same points of attack, has only a paralyzing effect. This means that there exists a close analogy between the ganglia and motor end-plate in so far as their pharmacological reaction is concerned. The explanation of this fact is that the cells of the end plate are derived from the cells of Schwann's sheath and are, therefore, neurogenic. Thus the basic difference in general structure assumed till now to exist between the peripheral neurons of the autonomic system and the spinal neurons ceases to exist.

Summarizing the results so far presented, I feel entitled to say that the neurochemical mechanism of the transmission of nervous impulses applies to all the synapses of the preganglionic and postganglionic fibers of the autonomic system and to the spinal nerves.

Beyond this conclusion there is missing only the proof that the chemical transmission from one neuron to the second holds true also in the central nervous system (C.N.S.). If this could be shown it would be established that synaptic transmission is chemical everywhere in the organism. Since this is a most important field it will be necessary to present the evidence for such an assumption.

TRANSMISSION WITHIN THE CENTRAL NERVOUS SYSTEM

As early as 1906 Sherrington suggested that the mechanism of transmission was similar at all synapses, and in 1925, after the discovery of the chemical transmission at the peripheral synapses, he discussed the possibility of this mechanism within the C.N.S. [29].

If acetylcholine is a transmitter at the central synapses it might be expected that it, as well as the cholinesterase, would be mainly located there at synaptic regions. As a matter of fact Nachmansohn [30] showed that a much greater concentration of esterase was found in the neuron-containing gray matter than in the white, and MacIntosh [31] later proved that the same condition was true for the distribution of acetylcholine. Most suggestive also are Nachmansohn's findings in regard to the relation existing between the concentration of esterase and functional development. In the brains of newborn rabbits and rats, which are markedly undeveloped, the enzyme concentration is very low and rapidly increases during the first three weeks after birth almost synchronously with the development of function in the brain, whereas in the brain of newborn guinea pigs, which are already developed at that time, the enzyme concentration is already almost as high as in adults. Furthermore Nachmansohn was able to show that there exists a kind of parallelism between the concentration of esterase and the functional development of different parts of the C.N.S. Barcroft and Barron have offered evidence for the early development of spinal reflexes and for a relative delay in the functioning of the brain in sheep embryos. Here, too, Nachmansohn found a correspondence with the appearance of esterase in the related parts of C.N.S.

From these findings alone we are, of course, not permitted to draw definite conclusions as to the part played by acetylcholine as a transmitter of nervous impulses within the C.N.S. We need complementary physiological evidence. This must be supplied from two main directions; it must be shown, first, whether or not acetylcholine acts like nervous stimulation upon the effector organs within the C.N.S. and, secondly, whether or not acetylcholine is liberated during stimulation of the C.N.S. as it is in the peripheral nervous system. There have been numerous investigations of the first point, some of which may be briefly indicated. The first experiments were by Dickshit (32), who injected small doses of acetylcholine into the cerebral ventrical of cats and found that the effect on respiration was similar to that produced by stimulation of the central end of the vagus. Quite a few investigators confirmed this result and demonstrated, in addition, a stimulation of certain hypothalamic regions. A great many scientists investigated the action of injected acetylcholine on the cortex of the human brain and of other animals, others applied the substance directly to the cortex, with the result that the electroencephalogram showed the same change as that produced by other kinds of stimulation of the cortex. In addition an increase of movements was often observed. Close intra-arterial injection was also effective. With regard to its incidental physiological importance it should be mentioned that the dose of acetylcholine producing an augmentation of the amplitude and an acceleration of the brain waves in some experiments amounted to only 0.1 to 0.2 gamma (one gamma = 1/1000 milligram), and the effect of these doses showed a close resemblance to the effect of electrical stimulation of the brain, on functional ability in the cortex.

From all these results there is no doubt that acetylcholine in adequate doses is able to increase activities going on in the different regions of the brain. This happens too with the transmitter in the periphery. It may be recalled in this connection that stimulation of the accelerator nerves increases the heart beat by releasing adrenalin and that stimulation of the vagus nerve increases the movements of the intestine by releasing acetylcholine.

Besides these effects consisting in modifying preëxistent activities, nervous stimulation (the effect of which is due to the release of transmitters or to the transmitters themselves) has in many cases an initiating action on some peripheral organs. The onset of muscle contraction or that of saliva secretion are examples. The same is true of the afferent nerves which start central reflexes. Of course the question arose as to whether acetylcholine acts in the same way within the C.N.S.; what should be expected if it is the transmitter.

Experimental evidence was provided that acetylcholine is able to initiate activity within the C.N.S. or, in other words, to provoke a primary discharge. The first experiments were performed by Cannon and Haimovici [33] in animals in which they had produced hypersensitivity by hemisection of the spinal cord with consecutive degeneration. By injecting acetylcholine intravenously in this state a primary discharge from the sensitized region was provoked, similar to that provoked by stimulation of afferent nerves. Quite recently Stavraky confirmed these results in other regions of the central nervous system. Previous sensitization, as induced in the experiments just quoted, is not needed for the demonstration of initiative action of acetylcholine within the C.N.S. Burn and Buelbring in experiments in which they only perfused the spinal cord succeeded, by injecting intraarterially very small doses of acetylcholine (one gamma), in provoking a primary twitch indistinguishable from the twitch produced by stimulation of sensory nerves. These experiments, taken together with those of Cannon and Haimovici and Stavraky, prove definitely that acetylcholine is able to provoke a primary discharge from motor elements within the C.N.S. exactly like that obtained by stimulation of nerves.

So far we have only considered the question as to whether acetylcholine is able to produce stimulating effects within the C.N.S. However, afferent impulses can have also inhibitory effects, namely, on reflex responses produced by other impulses acting at the same time. An instance is the reciprocal central inhibition of the antagonistic paired muscles. Though on the whole very little is known about the character of central inhibition, one point is definitely established, namely, that the impulse which, dependent on conditions, has a stimulating or inhibitory effect comes from the same nerve. Hence, it follows that the same transmitter must be responsible for both kinds of effects. In the peripheral nervous system we know quite a few cases where acetylcholine and adrenaline may have either a stimulating or inhibitory effect, depending on the state of the effector organ and other conditions. That the same holds true also for the C.N.S. has, to my knowledge, been proven thus far only in one case. Buelbring and Burn (34) have shown that the abolition of the knee jerk by nicotine is temporarily counteracted by acetylcholine. On the other hand, after the injection of eserine, acetylcholine depresses the knee jerk.

In summarizing the physiological results we may state that acetylcholine is able to augment or to inhibit activities going on in the C.N.S. and to initiate activities there exactly as it does in the peripheral elements. This taken together with the distribution of the cholinesterase and acetylcholine, as quoted above, makes it possible to say that, so far as we know, the latter acts as transmitter as it does in the peripheral nervous system.

In order to prove that acetylcholine is actually the transmitter it must be shown that it is released during central stimulation. Several investigators have sought such a proof: the first was Dickshit. He found, in previously eserinized animals, during stimulation of the central end of the vagus nerve, that acetylcholine was present in the cerebrospinal fluid and in the effluent blood. This result was confirmed by three other scientists, contradicted by two. Stimulation of the central end of the vagus induces great changes in circulation and respiration, which, in their turn, could create conditions for release of acetylcholine. Hence I do not regard the positive results so far obtained as definite proof. Minz and later on, Li dipped isolated spinal cords in eserinized Ringer solution and found that acetylcholine leaked into the solution during stimulation. Here, too, an objection can be raised: several investigators have definitely shown that excitation of isolated nerves provokes a release of acetylcholine from their endings. Since all the efferent and afferent nerves are cut by isolating the spinal cord the possibility cannot be excluded that the acetylcholine detected during stimulation originates from the cut nerves endings.

The question whether the addition of potassium provokes a release of acetylcholine, as it does in the periphery, has been investigated on isolated artificially perfused brains; an increase was found only with paralyzing doses. This, in my opinion, does not permit an unequivocal interpretation of the findings. Finally Buelbring and Burn (34) in their experiments with isolated cord circulation, stimulated the central end of the sciatic nerve and found acetylcholine in the effluent fluid. They are not certain, however, that this did not originate from the paravertebral muscles included in the circulation.

In summarizing we must conclude that the results so far obtained do not give a clear-cut answer to the question whether central stimulation releases acetylcholine at the synapses.

A great many experiments have been performed which approach the solution of this question in an indirect way. Within the peripheral nervous system under conditions where, for special reasons, it is impossible to demonstrate directly the release of acetylcholine, we are permitted to assume that it is released and transmits the impulse whenever eserine increases the effect of stimulation. Due to its anti-esterasic power eserine protects acetylcholine from being split, thereby permitting its accumulation. I must emphasize that, so far, it has not been proved that within the C.N.S. eserine acts only by preventing the destruction of acetylcholine or whether it exerts other actions in addition. This possibility has to be borne in mind because of the results of the following experiments.

Miller, Stavraky, and Woonton, [35] with direct application of eserine on the cortex of the cat's brain, observed very much the same changes of the electroencephalogram that are characteristic for stimulation by electricity or by acetylcholine. Quite often they observed, in addition, the onset of movements. Other investigators in perfusion experiments on the isolated cat's brain found that the addition of eserine produced increased reflex activity with spontaneous movements of the muscles of the head, and Buelbring and Burn, in their perfusion experiments on the dog's spinal cord, noticed on addition of eserine or prostigmine a discharge of impulses from the spinal cord recorded as contractions of limb muscles. Finally by applying eserine on the cortex in cases of petit mal an increase and an accumulation of attacks were noted.

How are these uniform results to be interpreted?

If we are not inclined to assume special actions of eserine within the C.N.S. which it does not possess in the peripheral nervous system, such, for example, as an ability to release acetylcholine or directly stimulate the C.N.S., the findings just quoted here have to be interpreted as meaning that acetylcholine is continuously produced and destroyed and that only by the anti-esterasic action of eserine is it accumulated to such a degree that excitation is produced. Be that as it may: even if this interpretation is accepted the experiments do not answer the main question as to whether acetylcholine is the transmitter. They only show that eserine at the spot where it is applied acts by itself or by its anti-esterasic power as a stimulus, whereby an excitation is produced which is propagated as is every locally produced excitation. Whether acetylcholine is involved in the synaptic transmission of impulses can only be determined by investigating whether or not eserine or other anti-esterasics influence the effect of reflectory stimulations.

There exists a great many careful experiments in this line. They were performed on different species of animals, with different procedures and anti-esterasics, with different doses applied in different ways, and with different reflexes investigated. All this may explain the variance in results. It is impossible to discuss the experiments in detail here. I must limit myself to considering the main result, which was that in all the experiments the anti-esterases, eserine and prostigmine, influenced the effect of reflectory stimulations. If we take the most likely view, that within the C.N.S. as in the periphery eserine and prostigmine act only by their anti-esterasic power, the results tend to show that acetylcholine is involved as transmitter in the outcome of reflexes.

The question whether or not acetylcholine is the synaptic transmitter within the C.N.S. seems not yet to be definitely solved. So far as I am aware no results have been obtained which would be contradictory to the assumption that acetylcholine is the transmitter; on the other hand, there exists, as we have seen, circumstantial evidence indicating the correctness of this view.

MECHANISM OF THE TRANSMITTER FUNCTION CONCLUSIONS

We may now examine the mechanism of the transmitter function. Obviously the site where the transmitters are released is the synapse, since the transmission takes place from the nerve endings to the effector organs. A synapse is not an anatomical conception. It is a purely functional one, meaning the place where the ending of the nerve fiber makes contact with the effector cell. Liberation of the chemical transmitter at the synapse naturally means its liberation from the nerve endings; otherwise it could not be the transmitter of the nervous impulse. That it really is liberated from the nerve endings was to be anticipated.

There is also direct and, in my opinion, conclusive evidence of the correctness of this assumption. It can, for instance be shown that after degeneration of the preganglionic sympathetic fibers the superior cervical ganglion loses almost all its acetylcholine content. Since the ganglion itself does not degenerate, the acetylcholine must have been derived from the endings of the preganglionic nerve fibers.

Here the question of the mode of liberation of the chemical transmitters arises. There are the following possibilities to be considered: (1) the transmitter, already present in nondiffusible form, may be rendered diffusible by nervous stimulation; (2) a diffusible transmitter may be newly formed following stimulation; or (3) a combination of (1) and (2) may occur. The first possibility involves the assumption that the transmitter is present in an inactive nondiffusible state. The fact that acetylcholine is found in the nerve ending is proof that it is present in nondiffusible form —or it would be destroyed by the esterase which is present everywhere. This is not merely an assumption; it is proved by the fact that so long as the physical structure of an organ is intact the acetylcholine content remains constant for hours. But acetylcholine that is added to the organ is quickly destroyed [15]. Another proof is the fact that no acetylcholine or adrenaline diffuses in the circulating fluid of organs without nervous stimulation. If the transmitter were diffusible the heart, for instance, would be continuously intoxicated, since 1/500 part of the amount of transmitter constantly present in it is sufficient to produce a maximal effect.

We have, therefore, to conclude that acetylcholine—or adrenaline, as the case may be—is present in the intact, nonstimulated organ in a labile, nondiffusible combination, in which it is inactive and not subject to attack by esterase. There seem to be a variety of such combinations. In the heart there must be a very unstable one, since the mere destruction of the physical integrity of the cells, for instance by grinding the organ, causes the acetylcholine to be quickly freed. Adrenaline, too, can be extracted by this means. Hence it seems probable that the nervous impulse acts by liberating transmitters from their nondiffusible combination.

Let us consider, however, the other possibility—that the impulse may act by provoking the new formation of a freely diffusible transmitter. There can be no doubt that the nerves are able to synthesize transmitters. Direct proof was given a few years ago by Quastel [16], who showed by *in vitro* experiments that an acetylcholine synthesis takes place in the isolated brain in the presence of oxygen and glucose. Recently MacIntosh [17] has proved the occurrence of acetylcholine synthesis in the nerve endings as well. He continuously stimulated the superior cervical ganglion, which was perfused with eserinized, oxygenated Ringer solution containing no glucose. The acetylcholine content of the ganglion and of the perfusion fluid and, consequently, the effect of nervous stimulation quickly decreased to a minimum. Continued stimulation after the addition of glucose to the perfusion fluid increased the release of acetylcholine almost immediately and restored the effectiveness of stimulation.

Quite recently our knowledge about the mechanism of the acetylcholine synthesis by nervous tissue has been advanced by a discovery of Nachmansohn [25]. He has found that in homogenized nervous tissue an enzyme, cholinacetylase, is present, which is able to synthesize acetylcholine in the absence of O_2 and glucose or pyruvate, provided that adenosinetriphosphate is present as a source of free energy.

Though these experiments prove that acetylcholine may be synthesized on nervous stimulation, I do not believe that the neurochemical transmission is effected by a primary new formation of acetylcholine in the nerve endings, for the following reason: Brown and Feldberg [19] have shown that the addition of small amounts of potassium to the perfusion fluid of a ganglion provokes a release of acetylcholine from the ganglion and its diffusion into the perfusion fluid just as preganglionic stimulation does. Since nobody will claim that potassium is able to cause a synthesis of acetylcholine, there remains only the possibility that it acts by liberating acetylcholine from storage in the nerve endings. There is no reason for assuming that nerve stimulation acts differently and synthesizes acetylcholine instead of liberating it.

We may conclude, therefore, that nerve stimulation causes the release of acetylcholine. Synthesis is a secondary process taking place after the disappearance of acetylcholine from the nerve endings.

The question that naturally arises at this point is: through what kind of a mechanism does the nervous stimulation liberate the chemical transmitter. It has been known for a long time that the propagated impulse within the nerves is connected with the shifting of electrical charges caused by the shifting of ions, especially of potassium, and we have learned, as just stated, that potassium liberates acetylcholine. Therefore it seems at least possible that the potassium liberated during the propagation of the impulse is in its turn responsible for the liberation of the transmitter during stimulation of the nerve.

Now that we have considered, however briefly, the essential factors concerned in the function of the nerve to liberate transmitters, the obvious question arises as to whether the nerve influences its effector organ by any other means. It has been shown that after application of substances antagonizing the effect of the transmitters, like atropine, ergotoxine, nicotine, and curare, the stimulation of the nerves concerned is completely ineffective. Therefore it cannot be assumed that the nerve influences the *specific* function of the effector organ by any means but by releasing transmitters. How could it, when the effect of the released transmitters is identical with the entire effect of nerve stimulation, even down to the smallest details?

CONCLUSIONS

The discovery of the neurochemical mechanism has finally provided a conclusive answer to the important, long-standing problem concerning the nature of the transmission of excitation from the nerve to its effector organ. The proof that this transmission is effected by substances liberated by the nervous system also furnishes the first evidence that the nervous system exerts specific chemical actions.

Of almost equal importance, it seems to me, is the disclosure of the nature of peripheral inhibition. Formerly it was impossible to conceive how stimulation of nerves could lead to an inhibition of activity in effector organs. Now that we know that this condition is brought about by release of chemical inhibitors, the matter is explained and appears, indeed, to be self-evident. Since we know that the process of excitation in the nerve is unspecific up to a certain point, there remains, I believe, no possibility whatever of imagining how nerve stimulation can inhibit an organ otherwise than by chemical means. In other words, the chemical mechanism is the only conceivable mechanism of peripheral inhibition.

The discovery of chemical transmission is important also from the point of view of general physiology. The fact that transmitters are liberated by the stimulation both of so-called augmentory nerves (whose effect is augmentation of function and of inhibitory nerves (whose effect is inhibition of function) proves that, whether the stimulation be augmentory or inhibitory in its *ultimate* effect on the end organ, the direct effect is always an increase in functional activity, so far as the direct point of attack, that is, the nerve ending, is concerned.

Knowledge in the field of pharmacology has been increased by the discovery and analysis of the neurochemical mechanism. Formerly it was quite generally believed that substances like the transmitters acetylcholine and adrenaline, as well as a great many other so-called vago- and sympathetico-mimetic substances, were effective by attacking and stimulating the nerves. There were two reasons for this assumption. The first was based on the fact that these substances have the same effect as the excitation of the respective nerves. Now we know that the reverse is true: nervous stimulation produces effects similar to those of the pharmacological agents in question, since these substances are the normal mediators of nervous impulses. The second reason was the manifestation of the action of antagonistic substances. Since the effect of the excitation of a nerve is nullified by the application of atropine, ergotoxine, nicotine, or curare it was concluded that these antagonistic substances paralyze the nerves, and, since they also counteract the effect of acetylcholine, adrenaline, etc., it was concluded that the latter stimulate the nerves. I have, however, been able to prove that atropine [20] and ergotamine [21] do not paralyze the respective nerves by showing that, after their application, nervous excitation is still effective in liberating the transmitters; the same condition was later shown to hold for the other antagonists, nicotine [22] and curare [23]. This result would obviously be impossible if the antagonists paralyzed the nerves, since the function of the nerves consists in liberating transmitters. Actually acetylcholine and adrenaline and the whole related group act directly on the effector organ, and the antagonists counteract this effect. The true point of attack of these important drugs has now been disclosed.

Finally, I should like to consider briefly the place of the neurochemical mechanism in the general picture of the relations existing between the cells. A characteristic of all living substance is its ability to react to stimuli. As our knowledge of the nature of stimuli grows, the evidence increases that stimuli of a chemical nature are predominant. A few years ago, for instance, no one would have imagined that mechanical stimuli may act indirectly by producing effective chemical agents. And now we have learned that the nervous system produces its effects by liberating chemical substances —effects until now looked upon as being due to stimuli of quite another nature.

Is this really surprising? Within the cell complexes of the organism which lack nervous elements, the relations between the cells can obviously be only of a chemical nature. Within those cell complexes having nervous elements, the nerve cell is not different in principle from other cells but differs only, in accordance with its special purpose, in being provided with processes, the nerve fibers. From this point of view it was, perhaps, only to be expected that the relations between the nervous system and other organs would prove to be qualitatively of the same nature as the relations between the other, non-nervous cells, that is, chemical.

For References see p. 311.

V

THE MATHEMATICAL NATURE OF PHYSICAL THEORIES

By GEORGE D. BIRKHOFF
Harvard University

THE prophetic conjecture that Nature is mathematical is one which goes back to Pythagoras and the ancient Greeks. The scientific development of the intervening twenty-five centuries has only served to establish this conjecture to a remarkable degree. The complementary fact that mathematics is *natural,* is however, just beginning to be grasped. On the objective side we have Nature, and, on the subjective side, we have its counterpart in mathematics, the language of Nature. These two aspects are as intimately related as are the two sides of a single coin.

To bring out the situation it is instructive to think of the development of science and language from the genetic point of view favored by those who think in biological-psychological terms. A child puts his hand too near the fire and is burned, and thereafter remembers that this A (touching fire) will bring about this B (pain and burn). The chain of association fixed in his memory is essentially of the propositional type "A implies B." He has learned a physical fact! Likewise in later situations he will find that not only does A imply B but B implies C, etc. The structure of thought is such that, by a kind of psychological shorthand, the intermediate term, B, is eliminated; and so he feels that "A implies C."

The essential genetic foundation here is obvious. The mental codification of the facts of Nature in logical and mathematical terms has its origin in the uniformity of Nature and of Mind. It becomes clear that, as John Dewey has recently said, "The basic importance of the serial relation in logic is rooted in the conditions of life itself" [1].

To carry the same idea further, let us note that each aspect of Nature which can be technically isolated has its own appropriate mathematical expression. Indeed, the main body of mathematics has arisen from three very simple aspects of the external world.

In the first place, our experience with events flowing in invariable sequence and recorded in memory leads to logic, as has already been indicated. It leads also to the associated notion of partially ordered classes or "lattices" [2]. These hierarchical arrangements of elements are found everywhere, but neverthless only recently has the theory of lattices been studied for its own sake.

Likewise the mental process of putting classes in one-to-one correspondence has as its external model a universe of distinguishable objects. The boy playing with piles of pebbles on the seashore is dealing with such a universe just as was the shepherd in ancient times who counted his flock by means of stones. In his ability to perform this simple process of one-to-one correspondence lies a basic difference between man and all other animals. It is self-evident that the matching game may be elaborated in various ways. For instance, two piles may be combined into a single pile. Practice in this technique led inevitably to the use of scratches and other marks, like the figure 3, for example, to designate the pile of reference containing three objects, with which all others might be compared. Thus the concept of number arises through experience with the universe of classes of identifiable objects, just as the concept of logical implication originates in the universe of ordered events.

A similar situation holds in regard to geometry. Here the universe from which we start is one containing idealized rigid bodies in ordinary space. These may be compared with one another by the process of direct superposition. Gradually in this way there spring up from experience with physical objects various ideas concerning geometrical relations. Eventually the ruler and the protractor are found to serve as particular rigid bodies with the aid of which it is especially convenient to make geometrical measurements.

Now it is well known that the vast body of pure mathematics deals principally with logic, number, and geometric form. Thus we may assert with reason that the whole structure of mathematics has been directly suggested by three exceedingly simple aspects or models of Nature, namely, the universe viewed as recognizable and invariable sequences of events, as a collection of classes of identifiable objects, and as a collection of rigid bodies.

With these three exemplars before them, mathematicians have set out to generalize and to modify. In the last century, the logical field has been greatly extended. Since early times other numbers such as zero, fractions, negative and imaginary numbers, and a host of other types have gradually been introduced as appropriate extensions of the integer. The motivation was not that afforded by any new model so much as that of generalization by abstraction, characteristic of mathematical thought. Likewise, after the systematic development by Euclid and others of ordinary geometry, there were invented n-dimensional geometries, non-Euclidean geometries, geometries in spaces of infinitely many dimensions and in abstract spaces.

To the modern mathematical mind, geometry appears as a less fundamental discipline than does logic or number. This means that one may dispense with the more elaborate model based on rigid bodies. The reduction of geometrical thought to its basis in number was accomplished largely by Descartes and Riemann. More recently the logicians, such as Frege, and Russell and Whitehead, have tried to discover number in logic itself, and thus to start from the model world of invariable sequences of events and nothing more. But the success of this ambitious and important venture remains doubtful, and number seems certain to hold forever a basic position in all mathematical thought.

The mathematician has created his grandiose mathematical structures, stimulated on the one hand by his own scientific imagination, and on the other by various suggestions and demands coming through aspects of the external world. His successes have truly been amazing.

In order to understand his method of attack and how it can be applied to physical theories, it is instructive to start with geometry as a convenient and familiar prototype. It will be recalled that Euclid makes use of certain "self-evident" truths, taken as undeniable and called axioms. In modern parlance these are called "postulates" and are thought of as hypotheses. The whole system of such hypotheses constitutes the "set of postulates" out of which the rest of geometry is to be developed logically as a body of "theorems." The general process of development has been twofold. On the one hand, the mathematician broadens the field of geometry by abandoning or modifying certain postulates in what seems to him to be an interesting manner. On the other hand, he spares no efforts in order to answer unsolved logical questions about geometry. For example, the Greeks propounded the question of "squaring the circle," and only after two thousand years of effort was this problem solved by Lindemann following in the footsteps of Hermite. It is not too much to say that if Archimedes were to return to us, a year or two of intense and carefully directed mathematical study would be necessary before he would be able to follow the argument involved.

The provisional and plastic character of sets of postulates was not fully understood until recently. The famous statement of Newton: "*Hypotheses non fingo*" (I do not frame hypotheses) indicates not merely that he had relied on his method of "Analysis" and subsequent "Synthesis," but also that he had not fixed conscious attention upon various presuppositions involved in his dynamical and astronomical theories. Similarly, his great rival, the mathematician and philosopher Leibniz, said:

> Far from approving the acceptance of doubtful principles, I would have people seek even the demonstration of the axioms of Euclid . . . . And when I am asked the means of knowing and examining innate principles I reply . . . that . . . we must try to reduce them to first principles, i.e., to axioms which are identical or immediate by means of definitions which are nothing but a distinct exposition of ideas.

An interesting and characteristic feature is that the postulates of Euclidean geometry appear in qualitative form. Such a postulate as that any distinct points A and B determine a straight line evidently does not essentially involve the notion of number. By successive steps it is then found possible to introduce quantitative ideas such, for instance, as are involved in the Pythagorean theorem. The final purely quantitative expression of Euclidean geometry is to be found in the analytic geometry of Descartes, which is entirely based on number. Similarly, it is possible to start with the analytic formulation of geometry, and quickly establish by an inverse process the validity of the qualitative postulates first mentioned. This kind of treatment of a mathematical theory may be indicated by the schema:

qualitative ←→ quantitative.

This double process is characteristic of mathematical thought in both the fields of pure and applied mathematics.

Someone not deeply acquainted with mathematics may well ask at this juncture: "But does this characteristic passage from qualitative to quantitative occur in the case of logic and number?" The answer is in the affirmative. It is not difficult to formulate a basis for the integers and other types of numbers in terms of qualitative postulates; indeed such sets deserve to be a matter of everyday knowledge [3]. For example, starting with a collection of objects which we may call integers and with an operation of "addition" of two of them, a, b, to give a third, which we may write $a + b$, a first natural postulate would take the form $a + b = b + a$. The answer in the case of logic is not quite so obvious and has only been recently recognized. Nearly a hundred years ago the Irish mathematician, Boole, saw that logic was a kind of abnormal algebra in which there were only two symbols of quantity, namely, 0 and 1. It has been realized only within a few years that logic permits of a normal algebraic formulation, provided that the fundamental operation of addition which Boole employed is modified. Thus, starting from a simple set of qualitative logical postulates, we obtain an *analytic* logic based on numbers in the form of ordinary sequence of 0's and 1's [4].

Before passing on to the consideration of other models of physical interest and their postulational treatment, it may be remarked parenthetically that, especially in the late nineteenth century, the method of advance in theoretical physics was sharply contrasted with that employed in mathematics. In fact, the mathematical physicist was inclined to try to deal with the results of physical experiment without concerning himself more than was necessary about mathematical questions. It is true that the grandiose successes of Newton were largely achieved through his remarkable mathematical genius, but after him it gradually became more and more the custom to look upon mathematics as merely a tool. As indicative of this attitude, I recall that not many years ago, after I had answered a query by a distinguished physicist, he said: "I wish I had a mathematical slave."

However, nowadays it is hardly possible to regard mathematics as only a useful handmaid, since for nearly four decades, the principal advances made in theoretical physics have resulted from mathematical insight and have involved a wider and wider circle of mathematical knowledge. In this period the successful theoretical physicists have been those endowed with mathematical vision, who believed in the power of the mathematical idea. The rank and file of physicists who lacked this quality have had to occupy themselves either with verifying the theories so obtained in the laboratory or with elaborating details in the new theories.

It is to be hoped that in the future more theoretical physicists will command a deep knowledge of mathematical principles; and also that mathematicians will no longer limit themselves so exclusively to the aesthetic development of mathematical abstractions.

We turn now to mention briefly two further simple physical models and their postulational forms. The first is that of a time series; and the second is that of a system of forces acting on a point. It may be remarked that the notion of time series has been the basis for an almost infinite amount of philosophic speculation; for example, a good deal of Bergson's theory of creative evolution is based on the metaphysical thesis that the theorem which results with postulates for time, like those given below, is somehow not valid.

The postulates for the time series may be formulated readily: The "undefined class" of objects which we consider is that of "instants of time"; the "undefined relation" between two (distinct) instants is that of "precedence"—thus if A precedes B in time, we write $A < B$.

A set of qualitative postulates is then the following:

I. Either $A < B$, or $B < A$, but not both, are true. If $A < B$ and $B < C$, then $A < C$.

II. For any A there exist X, Y such that $X < A < Y$.

Definition: An (open) interval is a set of points X of one of the four types: $A < X < B$; $X < A$; $B < X$; X arbitrary.

III. If $---)A_{-1} < A_0 < A_1 < A_2 ---$, then the set of A_n and of the intervals $(A_n A_{n+1})$ constitutes an interval.

IV. Any two intervals can be put in $1-1$ correspondence, preserving order ($<$).

V. Any set of distinct intervals is numerable.

On this basis it is possible to prove that with each instant of time there may be associated a numerical variable which ranges from negative infinity to positive infinity, in such a way that if $A < B$, the number of A is less than that of B, and vice versa [5]. This end-result or theorem constitutes the fundamental quantitative theorem of the theory. It is obvious, conversely, that if we have a collection of objects associated with numbers in this way, and interpret $A < B$ as meaning that the number of A is less than the number of B, then all of the preceding system of postulates will be fulfilled.

The great English mathematical physicist, James Clerk Maxwell, sometimes found it amusing to put his ideas in the form of mathematical verse. I have ventured to attempt something similar with the above set of postulates for the time series and the fundamental theorem to which it leads; to this my friend, the poet David McCord, has given, at my request, a more poetic form:

The instants of time form a limitless sequence.
Precedence still rules them no matter their frequence.
Each infinite set in an interval settles,
Resembling its kin as a petal does petals;
Besides, one can always in sequence review
A collection of intervals standing in queue.

This granted, good reason not rime will define
That the instants are ordered like points on a line.

We turn next to our first obviously physical model, namely, that of a set of forces acting at a point. Almost every person has an intuitive grasp of the idea of forces, and their combination in a resultant force. A possible qualitative set of postulates for this simple physical universe is the following:

1) Forces which act in the same line have a resultant whose magnitude is the sum of the magnitude of the component forces.

2) The resultant of forces does not depend on the order or grouping of the constituent forces, varying continuously with them.

3) The resultant of forces is independent of the unit and of orientation in space. (Principle of sufficient reason.)

It will be observed that here the postulates are not wholly qualitative, since the idea of the magnitude of a force is involved. In all physical theories the idea of number is accepted for convenience as if it were qualitative.

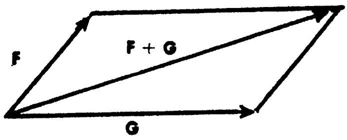


Fig. 44. Illustrating quantitative result of the Parallelogram Law.

The final quantitative result in this case is the familiar parallelogram law embodied in Figure 44.

In this case David McCord's verse is as follows:

The resultant of forces in line
Is the sum of these forces. Combine
Two or more which don't happen to act
In one line, it remains as a fact
That their order and grouping don't count.

And again, disregard the amount
The resultant is thrown out of joint
Should some forces be changed half a point.
Change a unit, direction in space—
Your resultant still falls into place.
When forces so act without flaw,
They obey Parellelogram Law.

It will be noted that, in the last postulate, reference is made parenthetically to the principle of sufficient reason of Leibniz. This is an extremely useful principle which may be employed whenever there is an underlying ambiguity. In the present case the ambiguity is that of orientation in space and of the unit of length. Later on, we shall observe that such ambiguity is always associated with a "group."

Let us attempt the direct use of the principle before such further elaboration. In the first place, let us observe that if there are two equal forces F, their resultant force must lie in the plane which they define. For why should it fall on one side rather than on the other? Likewise it is clear that this resultant must lie along the bisector of the angle formed by the two given forces, for a similar reason. These simple conclusions indicate how the principle may be used.

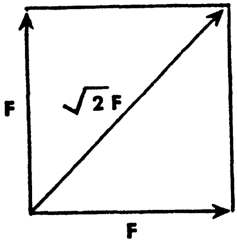


FIG. 45. Illustrating resultant of two equal forces acting at right angles.

As a matter of fact we propose to go one step further and show that, in the special case when the two equal forces F are at right angles to one another, their resultant must be precisely the diagonal of the square which they bound, as required by the parallelogram law, as shown in Figure 45. This means that the magnitude of the resultant force must be F multiplied by the square root of two ($\sqrt{2}$ F), according to the Pythagorean theorem. Of course we already know that the resultant will fall along the bisector in virtue of what has just been said above.

In order to prove our statement we start with the given force F and add to it two equal forces of one-half the magnitude, F/2, which are at right angles to the given force, see (b) in Figure 46. Moreover we break up the force F into two parts also of magnitude F/2, one associated with each of the forces first introduced. At this stage, see (c), we have a system of two pairs of forces whose entirety must be equivalent to the force from which we started, on the basis of the postulates written above. But now consider one of these pairs of two forces of magnitude F/2 and mutually perpendicular. Their resultant will fall along their bisector with a magnitude k times the magnitude of the two components. Thus the two pairs of forces reduce to two forces of magnitude kF/2 at right angles to one another, see (d). Finally, since the size of the unit does not matter, according to the principle, we see that the final resultant will have a magnitude k times that of each of the two components, namely, $k^2F/2$, see (e). But this resultant force must, of course, be nothing other than the force from which we started. This shows that k *must be* the square root of 2.

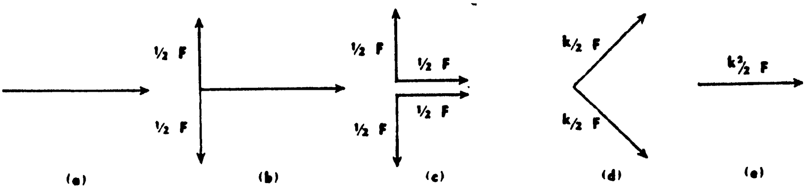


FIG. 46. Illustrating mathematical proof of the Parallelogram Law.

The short argument just made really constitutes a genuine introduction to the mathematical proof of the parallelogram law. It was this type of proof, for instance, which was given by Laplace in his *Mécanique Céleste*.

With these preliminaries in mind, we are in a position to consider advantageously physical models of a less elementary type and, in particular, such as satisfy the basic principle of causation or uniformity. This principle asserts that a given initial state of a system determines the unfoldment of all of the subsequent states.

Between the many theories of this causational type there is a difference in character. Some have been designed to account for a special aspect of nature, but were never intended to be taken as foundational for physics as a whole. We will call the models used in such cases *proximate* models. Other models have seemed to contain within them some germ of possibility of universal application, and these will be termed *ultimate* models. Before taking up the study of some ultimate models, it will be worth while to consider briefly one important proximate model, namely, that involved in the so-called analytic theory of heat.

Suppose that we consider a set of rigid bodies fixed in position in ordinary space. Each of these is thought of as isotropic and characterized by a certain ability to conduct heat and to contain it. In other words, at each point of these bodies there is a certain specific conductivity and a specific heat, and these constants are taken to be the same everywhere in one and the same body. The fundamental independent variables which enter are those of position and time, whereas the fundamental dependent variable is the temperature. For this theory the postulates are usually stated in simple quantitative form, although they readily admit of partially qualitative formulation, as when it is said that heat always flows from warm to cold. These postulates lead to a single central differential equation, the so-called equation for the flow of heat. Thus the definitive quantitative form of the theory is contained in this equation. By its aid we are able to determine the flow of heat which will result from a given initial state of the system. In this way Fourier was able to solve a great variety of problems. It is interesting to recall how Kelvin later on applied Fourier's theory to the problem of estimating the age of the earth. For this particular case the earth was taken as an isotropic sphere at high initial temperature from whose surface heat was conducted away into empty space according to known physical laws. By using the known data concerning the temperature in deep mines, it was possible to obtain a reasonable quantitative estimate. Of course since the discovery of the phenomenon of radioactivity this estimate has been discarded. Evidently this model of Fourier is one of proximate type.

Later on, heat was explained by Maxwell and others by means of the so-called kinetic theory of heat. In this theory heat is conceived of as a mode of motion of the atoms constituting the body in question. Roughly speaking, a rise in temperature is taken to correspond to a greater average velocity of the atomic particles. With this interpretation, it is possible to justify the postulates of the analytic theory of heat. Evidently the kinetic theory is of more basic type and might provide an ultimate explanation of heat—if the universe were dynamical instead of electrodynamical!

The conceptual models which have been employed by physicists are extremely varied and numerous. It almost seems as if Nature had desired to anticipate every possible mathematical system and to employ every kind of invention. My distinguished friend and colleague, Bridgman, to whom I made a remark of this sort some time ago, replied that, so far as he knew, Nature had never used the principle of the wheel! At any rate, the varied models exemplified in Nature seem to be beyond the range of human imagination, and we are only commencing to understand a few of them.

Practically all of the proximate models of Nature share certain characteristics due to the fact that they accept the principle of uniformity. The postulates, expressed in terms of number, soon lead to certain basic differential equations which are expressed in the symbolism of the calculus. Thus the theory of the model is reduced to the solution of differential equations. The treatment of the motion of our planetary system provides a well-known example.

We now pass to a consideration of the principal attempts at ultimate theories from the time of Newton onward. So far as their mathematical nature is concerned, we may note in advance the following general conclusions: There is in each case an underlying four-dimensional conceptual framework of space and time, with the possible exception of the latest quantum-mechanical phase. In each case there is a fundamental ambiguity and corresponding mathematical "group" so that the principle of sufficient reason may be applied. Finally, corresponding to each group there is an appropriate mathematical language.

To a considerable extent the various types of theory may be obtained from one another by a process akin to transposition to the modified language.

We shall also attempt to point out some of the mathematical difficulties and physical limitations in the various cases. In conclusion, some remarks of a general character will be made.

Classical physics starts with the framework of Absolute Time and Absolute Space. This is the framework to which all of our common-sense experience in everyday life inevitably leads us. Let us consider the character of this framework from what may be called the group-theoretic point of view.

With Newton, we think of time as an equably flowing variable, the same for all observers. However, there is an underlying ambiguity in that the various instants of time are wholly indistinguishable from one another. Once, however, a particular instant or epoch has been chosen from which to measure time, all other instants may be characterized uniquely by their temporal distance from the epoch. Thus the epoch chosen for practical purposes is the beginning of the Christian era.

It is interesting to recall that, from the philosophic point of view of Leibniz, this ambiguity of structure precluded the possibility that Time is an Absolute Being. Briefly, his argument was that God would reject Time because none of the decisions of God are ambiguous. In this way Leibniz concluded that both Space and Time were relative rather than absolute.

For us, however, it is the group involved, reflecting this ambiguity, that is significant. To explain this idea, let us fix attention upon a square ABCD. There are four operations of rotation which leave the square in the same position: namely, the operation I (the so-called identity) which does nothing; the operation R which rotates the square through 90°; the operation S which rotates it through 180°; and the operation T which rotates it through 270°. The successive combination of any two of these operations amounts to a third. Thus R, followed by S, yields T: RS = T. In this way we get a "multiplication table" of the group of operations which embodies all of its essential properties. In Figure 48 we give this table. It will be noted that for each operation there is another which undoes what the first does. Thus R followed by T does nothing, that is RT = I.

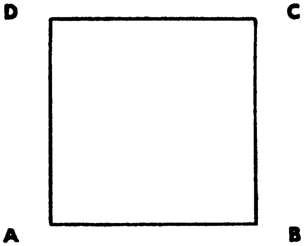


Fig. 47. Operations of rotation with square ABCD.

	I	R	S	T
I	I	R	S	T
R	R	S	T	I
S	S	T	I	R
T	T	I	R	S

Fig. 48. Diagram showing group operations of square (Fig. 47).

Any collection of operations which can be combined in this way and which includes, along with any operation, another inverse operation which undoes what the first does, is called a group. The identical operation, I, which does nothing, is always considered to be a member of the group.

In the case of the model of Absolute Time, the operations of the group are those which add or subtract a constant value to or from the measure of time, or, what is really the same thing, shift the epoch. It will be observed that such a shift will not change in the least the interval of time between two instants. Such an unchanging quantity under the operations of a group is said to be an "invariant." Likewise, we may speak of "invariable properties." Thus in the case of the group of the square, the property that two vertices are opposite or adjacent is invariable.

In Absolute Space there is a similar ambiguity but of more extensive type. All directions in space and all reference points are on a common basis. It was for this reason that

Leibniz asserted that Space was not an Absolute Being. From the point of view of the theory of groups, we may say that all translations and rotations of the space into itself form its group, and the corresponding invariant is that of the distance between two points. There are also other invariants, such as the angles between two lines, but these may be considered to be derived from the still more basic invariant of distance.

It could be shown mathematically that the theory of one-dimensional time and three-dimensional space, as presented ordinarily, is contained in the theory of the associated groups and their invariants.

Now, although Newton did not explicitly recognize the fact, there was involved in his theory a more extensive ambiguity than that to which we have referred above. His underlying postulates were such that they did not distinguish between any particular framework and any other which was moving with uniform velocity in a straight line with respect to the given framework. Furthermore, there is almost nothing in the Newtonian theory which differentiates between units of length or units of time. In this classical theory there appear two fundamental invariants, namely, the interval of time between events, and the distance between points at a certain time.

Mathematics has provided a classical language which is appropriate to the Newtonian theories and the underlying group. This is the language of 3-vectors. A 3-vector is a directed quantity, like a velocity, which is characterized by its direction and magnitude in ordinary three-dimensional space. This abbreviated vectorial language is used constantly in all of classical physics when it is desired to express the theory in as condensed terms as possible.

As I have shown elsewhere, the essential significance of Leibniz' principle of sufficient reason is group-theoretic [6]. The general relationship involved may be indicated by the synoptic formula:

Metaphysics ⟷ Principle of Sufficient Reason ⟷ Theory of Ambiguity and Groups

Of course from our present point of view, we must think of physics as a special branch of metaphysics.

To bring out more clearly the application of this group-theoretic principle, let us consider for a moment the law of gravitation of Newton. If we have two gravitating particles, it is clear that the forces between them must lie in the line joining the two particles, inasmuch as spatial rotations about this line are operations of the group of motions, and the resultant forces must be independent of these operations. Likewise if the two particles are relatively at rest, the forces can only depend upon the mutual distance. Of course it is natural to take the two forces to be equal, since observation shows that under ordinary circumstances action and reaction are equal; and it is *simplest* to suppose that the force only depends upon the distance even if the particles are in relative motion. Furthermore, it is philosophically plausible to conjecture that the force varies continuously with the distance and tends to become small at great distances. In this way the law of gravitation of Newton is indicated as one of the simplest and most reasonable possibilities, on the basis of the underlying group and the requirement of mathematical simplicity. Another equally simple possibility is that the force might be inversely proportional to the first power of the distance instead of the second power as in the Newtonian law. But this possibility is ruled out by Kepler's planetary laws for the motion of two attracting bodies.

Nearly all of classical physics of the dynamical era has been constructed more or less consciously on the basis of group-theoretic considerations mixed with philosophical requirements of the type which we have just indicated. This method has shown itself enormously successful in advancing our knowledge of the physical universe.

Besides the conceptual framework of space and time which forms the stage of classical physics, we must specify conceptual models of matter which may occupy this space in time. The starting point is naturally the particle and the rigid body. Then follow models convenient for hydrodynamics and elasticity, such as isotropic adiabatic fluids and isotropic elastic solids. Upon these may be imposed certain geometrical constraints, as, for example, when the center of a rigid sphere is constrained to be at a fixed point. More generally, various forces of attraction or repulsion consistent with the conservation of energy are allowed. Thus we are led to dynamical systems essentially characterized by differential equations of the Lagrangian type. These equations may be written down as soon as the explicit expressions of the "kinetic" and "potential" energies have been determined. The celebrated principle of the conservation of energy affirms that the total energy, both kinetic and potential, always remains constant.

It is a remarkable mathematical fact that the Lagrangian equations of a complicated dynamical system may be derived from such a unitary principle as the principle of least action. It was Maupertuis and Leibniz who first stressed the importance of minimum or variational principles such as that of least action in physics. Later on, we shall attempt to assess to some extent their actual importance for physics. At the moment, however, we merely observe that in the Lagrangian formulation it was the *difference* between kinetic and potential energy which was involved in this formulation, and this difference has no direct physical significance [7]. Similarly, there was an artificial element in the Hamiltonian variational condition, called Hamilton's principle, which was later obtained and which apparently involved only the total energy H, for, in order to express the equations of motion by means of this new principle, it was necessary to double the number of dependent variables.

Despite the artificiality of such variational principles, they were regarded as of outstanding theoretical significance throughout the classical era. The dynamical formulation of physical laws was regarded as likely to lead to an ultimate explanation, even after it was recognized that electricity played an absolutely fundamental role in the atomic domain. In fact, Maxwell found that his electromagnetic equations could be derived from a suitable variational principle, and it became the consensus of opinion that in such principles was to be found the mystical element uniting matter, electricity, and the ether.

It was not noticed that such an ether did away completely with the relativity of motion which had been present in the Newtonian framework. The group allowed was now reduced to a mere change of point of reference or of orientation in space.

From the new point of view, rigid bodies and other types of matter were thought of as the bearers of electric charges whose motions were to be determined not only by the physical forces of dynamical origin, but by the electrodynamic forces as well. In this spirit, suggestive models of atomic structure were proposed by Rutherford (1911) and others, on the basis of a positively charged nucleus surrounded by a certain number of negatively charged electrons in positions of equilibrium. Furthermore, a great many optical phenomena, such as optical rotation, double refraction in crystalline media, etc., were accounted for. To be sure, in each case an *ad hoc* atomic model was constructed in order to obtain an explanation.

Thus at the end of this dynamical era, there was prevalent in most physical quarters a conviction that by a judicious combination of dynamical and electrodynamical laws governing neutral and electrically charged masses moving in a fixed electromagnetic ether, it would be possible to furnish an explanation of all the phenomena of Nature. The principle of least action and other variational principles were conceived as somehow affording the basis of unification.

The purely electromagnetic era was ushered in by the remarkable experiments of Michelson and Morley in 1887. Taken at their face value, these suggested that matter was not in the least affected by its motion through the ether, and it became natural to extend this idea as far as it could be conveniently applied. Around 1892 the Irish mathematical physicist, Fitzgerald, and his Dutch contemporary, Lorentz, made the hypothesis that matter was contracted in the direction of its motion by a certain fraction dependent upon its velocity with respect to the ether. In this way the negative results of the Michelson-Morley experiments could be adequately accounted for. Somewhat later another Irish mathematical physicist, Larmor, assumed explicitly that a similar contraction took place in time, that is to say, moving clocks were supposed to be slowed down in the same ratio. In this way, it became gradually apparent that the complete electromagnetic equations could be treated so as to afford a natural explanation of the null effect of motion through the ether.

Today it is evident that, by the thoroughgoing use of the "Lorentz group" associated with this type of spatiotemporal contraction, it is possible to give a qualitative approach to the electromagnetic equations so that they appear to be as natural from the group-theoretic point of view as does Euclidean geometry. It must be emphasized that, before this group and its associated ambiguity were introduced, the exact form of the electromagnetic equations had appeared as extraordinarily artificial. Here, then, was a first great triumph of the purely electromagnetic point of view.

Another remarkable fact is that the form of relativity of motion appearing in the new theory is really that which would be suggested naturally to an astronomer who looked out upon the stellar universe with completely impartial view. To him all of the stars would seem to be on an equal basis, and purely mathematical argumentation, in which it was postulated that such an equivalence existed, inevitably leads to the same form of space-time framework as before. We may therefore call the space-time background "astro-electric."

This modification of cosmic outlook was brought about largely through the work of Larmor and Lorentz, but it was Einstein who first, in 1905, embodied it in his "special theory of relativity."

In the new world thus set up, there is essentially only one invariant, which Lorentz called "local time"; it will be recalled that in the older theory there were the two invariants of distance and of time. The group involved restores the relativity of motion which had been lost when the concept of a stationary ether made its way into physical speculation. This

Lorentz group mixes up space and time inextricably so that the concept of simultaneity no longer retains absolute significance. Nevertheless, the new group involves exactly as many arbitrary constants (ten) as the Newtonian group, and hence it is fair in a certain sense to say that the new relativistic framework is no less absolute than the classical framework!

At the present time, every worker in the electromagnetic theory assumes that the background in which phenomena take place is of this type. The reason is that, in this way, a complete mastery of the equations and their transformations is readily obtained.

The basic language now becomes that of 4-vectors, whereas in the previous Newtonian era that of 3-vectors had been fundamental. The new language obeys the same formal laws as the old, but the vectors involved have four components instead of three.

What is lost in this change is the possibility of an analogue of the rigid or elastic body. Since an electrically charged body tends to expand, due to the enormous forces of electrical repulsion which exist, something like the rigid or elastic body is necessary to prevent the atom from exploding. However, it has not been found possible to devise any adequate substitute for such bodies in electromagnetic space-time, so that this difficulty remains to the present day. It might be thought that the point particle of electrically charged matter would provide an adequate carrier of the electrical charge, but, due to the fact that infinite forces and energy are involved, the mathematical difficulties of dealing with such particles are found to be extremely great, and it seems doubtful that it is possible to employ the particle model as a basis.

It thus becomes a fundamental question as to what types of matter can be used. Cosmic dust, that is, matter characterized by density without pressure, can collapse to a point of finite mass, can interpenetrate other cosmic dust, and can even turn inside out. Furthermore, electrically charged cosmic dust would tend to expand and scatter with incredible rapidity. If, however, we introduce the simplest type of matter in which pressure occurs, namely, an isotropic adiabatic fluid, there is a fundamental difficulty which arises when two portions of fluid collide from opposite directions at velocities exceeding the disturbance velocity in the fluid. In this event the mathematical equations themselves break down.

Consequently it appears that there is really only one type of matter which is self-consistent from the mathematical point of view, namely, that in which the disturbance velocity at all densities is precisely the velocity of light. In 1926 I introduced this form of matter under the name of the "perfect fluid." The equation of state of such a fluid was found to be simply (in units of seconds and light-seconds):

$$\text{pressure} = \tfrac{1}{2}\ \text{density}.$$

This fluid is supposed to be in equilibrium at a "cosmic pressure" not zero. The perfect fluid is very nearly incompressible and of almost invariable mass under ordinary circumstances.

It should be added that up to the present time there has been no indication whatsoever that physicists would be at all interested in such a perfect fluid just because it is the unique form of isotropic adiabatic fluid which can carry electricity without introducing mathmatical difficulties. Indeed, the question of mathematical consistency has never worried theoretical physicists very much.

It is my considered opinion that isotropic adiabatic fluids other than the perfect fluid are not only faulty in mathematical construction, but are aesthetically inappropriate to the background of electromagnetic space-time. Nevertheless, only such forms of isotropic adiabatic fluid have been used by Einstein and others for theoretical purposes.

In 1912 Nordström attempted to develop a simple theory of gravitation which would be appropriate to the new space-time. His theory accounts adequately for ordinary gravitational phenomena. However, it was at once noted that, according to this theory, the perihelion of Mercury would experience a retrogression rather than the advance which had been observed. Consequently this theory was immediately abandoned.

Very recently [8] I have found that by introducing the perfect fluid as the basic form of matter, it is possible to formulate in the same language of 4-vectors a simple theory of gravitation which accounts for general gravitational phenomena, and also explains the three crucial effects that have been observed. One of these effects is that of the advance of the perihelion of Mercury referred to above, and the other two are the bending of light by the Sun, and the shift of spectral lines toward the red in light coming from the Sun, both of which had been predicted by Einstein's theory of 1916. My theory was suggested in part by the form of the theory of Einstein, and part by the Nordström theory. All of these attempts by Nordström, Einstein, and myself are in a certain sense direct transpositions of the theory of Newton, on the basis of 4-vectors instead of 3-vectors in the cases of Nordström and myself, and on the basis of the tensors appropriate to curved space-time in the case of Einstein.

In my opinion, the conceptual possibilities of the electromagnetic framework deserve the closest study. Unfortunately, changes in the point of view of theoretical physicists have been so rapid in the last decades that promising directions of advance have been hastily given up without due exploration of their possibilities. The mathematicians, if not the physicists, should take up the highly important task of studying more deeply the construction of physical models on the basis of electromagnetic space-time.

When it was found that Nordström's and allied attempts at direct extension of the Newtonian gravitational theory to electromagnetic space-time failed to provide an explanation of the first crucial effect specified above, it was natural to try to devise other forms of gravitational theory. The outcome was Einstein's theory of relativity of 1916 which seemed, at first glance, to suggest that physics might be nothing more than a kind of four-dimensional geometry.

This theory was based on the hypothesis of a curved space-time in which the small portions of space-time are essentially of the same type as electromagnetic space-time. The single invariant of local time (analogous to arc length) was made to depend upon ten gravitational potentials, and the theory came under the type of four-dimensional geometry studied earlier by Riemann. Ricci and Levi-Civita had shown, too, that for the expression of geometrical properties in such a Riemannian geometry, the language of tensors was appropriate. The properties of such tensors were developed in the "absolute differential calculus." Fortunately, Einstein found this instrument available in the elaboration of his basic principle of equivalence. This principle affirmed that space-time is curved by matter, and that the ultimate parts of matter follow straight lines in this curved space-time so far as is possible.

Now it was postulated at the outset in the new theory that the underlying "group" was general, so that there was no longer any kind of absolute underlying framework. In other words, proper independent variables were no longer available. This fact enormously increased the mathematical difficulties of the theory without yielding counterbalancing advantages in any other direction. For this reason, the generalized relativity of Einstein has never entered effectively into physical speculation. Larmor probably expressed the general opinion among physicists when he called this theory "an auxiliary construct," and said "the absence of space and time and motion in the auxiliary construct is against reality" [9].

One advantage of the alternative theory which I have mentioned is that it preserves the basic electromagnetic framework. Of course it is too early to say whether these theories or some other gravitational theory lies in the direct line of progress.

What was most effective, however, in preventing the new relativistic theories from being followed out thoroughly was the increasing realization that atomic phenomena seem to be inexplicable on any basis such as that provided by the background of a four-dimensional space-time continuum. It was Planck who, in 1900, introduced the revolutionary idea of quanta of energy in order to meet some of the difficulties. The later attempt of Niels Bohr (1913) provided a quasi-dynamical model of the atom which served to explain in a very suggestive way many spectroscopic facts on the basis of quanta of energy. It is true that Bohr, in achieving his purpose, introduced *ad hoc* hypotheses, but the fundamental advance was that he had provided a formalism by which predictions could be made. At the hands of Sommerfeld and other physicists who followed Bohr, many triumphs were scored by the theoretical physicists, and all turned to the study of atomic phenomena to which the new theories might be applied.

In these quantum-mechanical theories the hitherto basic concept of a four-dimensional continuum of space and time is completely abandoned without a qualm. Attention is generally fixed on a particular atom or molecule and its immediate neighborhood, and nothing else is considered.

The culmination of this type of theory came in 1926 when Schrödinger formulated his celebrated "wave equation." It is interesting to recall the general character of Schrödinger's treatment. He starts out with the conventional classical model of an atom as composed of a positively charged proton having a given mass, and corresponding negatively charged electrons which are of much smaller mass. The equations of motion are then of usual Hamiltonian form. Using certain vague optical ideas, Schrödinger finds a way to associate a certain linear wave equation with these dynamical equations. Under appropriate "boundary conditions," this Schrödinger wave equation has solutions possessing certain frequencies and these lead to the determination of the position of atomic spectral lines, of their intensities, etc. Looked at in the cold light of logic the processes involved seem to be a kind of mathematical abracadabra, but the result is extremely successful and suggestive just the same. The lack of cogent reasoning throughout has often been justified by the assertion that light and electromagnetic phenomena are to be explained alternatively on both a particle and a wave basis!

To give a correct impression of the significance of such wave-mechanical theories, it is necessary to make several comments. In the first place, it had always been supposed that light emitted from an atom was the result of electromagnetic vibration, and it was reasonable to expect that the equation for such a vibration might be a linear partial differential equation, like the Schrödinger equation. But no conceptual method of obtaining his equation has been advanced except my own very incomplete attempt of 1926 [10] when I introduced the concept of the perfect fluid. In this theory I supposed that the proton and electron could freely interpenetrate so that in the normal state of equilibrium there was complete neutralization of the electric charge. Under disturbances of the state of equilibrium, certain frequencies of vibration were obtained. This theory seems to me to be obviously in need of further development. Perhaps it does show, however, that an explanation based on a four-dimensional space-time continuum of electromagnetic type is within the bounds of possibility.

To my mind, the Schrödinger equation is a successful attempt to find underlying equations, without providing a suitable conceptual setting. My notion can be illustrated by the following situation which might have occurred in the development of Newtonian gravitation. It will be recalled that Kepler formulated three simple and well-known laws governing the motion of two bodies, such as the Sun and the Earth. Now suppose that someone had propounded the following theory: In the planetary system no other laws operate than these laws of Kepler between two members of the system. From time to time the pair of bodies moving according to these laws changes arbitrarily, just as Newton had conjectured his light corpuscles might have "fits of easy transmission and reflection." Meanwhile, all of the other bodies not involved in the binary Keplerian dance are supposed to move in a straight line at constant velocities. It can then be shown that the motion of the bodies in the planetary system must be essentially along Newtonian lines [11]. Furthermore, there is evidently involved an indeterminacy principle very analogous to the indeterminacy principle introduced by Heisenberg and regarded as basic in quantum mechanics. Or shall we rather say that the change of partners in the binary dance is governed by an exclusion principle like that of Pauli?

Yet who would declare that such an *ad hoc* theory of the planetery system, which would be nearly equivalent to the Newtonian theory, as far as predictive power is concerned, is an adequate substitute for the grandiose conceptual theory of Newton?

It seems to be inevitable that somehow or other a background of space and time must eventually play a part in any truly satisfactory account of our physical universe.

Having completed this hasty analysis of the development of basic physical theory since the time of Newton, we are in a position to summarize our general conclusions advantageously.

In the first place, all of these theories, except that of quantum mechanics, employ a selected four-dimensional framework of space and time which is to serve as the stage for the interaction of matter and electricity. In all of these, the principle of the uniformity of Nature is taken for granted. The laws which are postulated generally lead to one or more ordinary or partial differential equations, so that finally the basis of all theory turns out to be a set of differential equations—the differential equations of the physical universe!

Whether or not quantum mechanics, in which space and time only function as preliminary scaffolding, can be accommodated within such a framework remains to be seen. However, the illustration given above may serve to show how it is possible to prefigure the formulas of a theory without providing a suitable conceptual basis. Thereby it is suggested that the fundamental Schrödinger equation may amount to a genial formalistic conjecture, yet to be justified.

It was also apparent that in each theory there was a basic underlying group, with corresponding invariants and mathematical language. In the classical case the group was that of Absolute Space and Time, augmented by uniform translatory motions; distance and interval were the two fundamental invariants; and the basic language was that of 3-vectors. Here the underlying model was dynamical, and the space and time were those of everyday intuition. Furthermore the accepted laws were always chosen in accordance with the underlying group and the allied principle of sufficient reason. The dynamical formulation was extremely elegant in terms of suitable variational principles, such as the principle of least action and Hamilton's principle.

As electromagnetism began to be developed, the concept of the ether was introduced; this meant a limitation of the underlying group, since motion now appeared as absolute in the ether. The so-called electromagnetic equations could be given 3-vector form, but their formulation seemed entirely artificial, in marked contrast to the equations of the older Newtonian physics, for example, those embodied in the so-called parallelogram law for the composition of forces.

But a little before the beginning of the present century, when the electromagnetic nature of the atom began to be appreciated and when the Michelson-Morley experiments and others had indicated the complete indifference of matter to its passage through the ether, there began to occur a fundamental change in outlook. This eventuated in the use of a new fundamental four-dimensional substratum of space-time, that of the special theory of relativity. The model here used was "astro-electric," in that it was equally appropriate from the point of view of electromagnetic and of stellar phenomena. The underlying group was the Lorentz group; the single invariant was local time; and the appropriate language was that of 4-vectors.

The necessary reformulation of the fundamental properties of matter and electricity amounted to a modulation of the older theory in the direction suggested by the change in the basic mathematical language. Here there was the great advantage that, on the basis of the principle of sufficient reason, the electromagnetic equations appear natural against the new spatiotemporal background. But counterbalancing disadvantages appeared, in particular the convenient concepts of the rigid and elastic body had to be abandoned. The theory of gravitation of Nordström (1912) was based on the new framework but was at once given up, since it did not predict the observed perihelial advance of Mercury. However it has been pointed out that a new simple theory (based on the "perfect fluid") is now available in this type of space-time.

With this electromagnetic framework the idea of energy became more obscure, and variational principles seemed to retain only limited significance. Moreover no way was found of improving the older treatment of optical and electrical phenomena.

Einstein's theory of gravitation of 1916 also had its group (the general group), its one invariant (local time), and corresponding language (that of tensors). It was based on the model of a space-time curved by matter, and afforded a brilliant account of gravitational phenomena. Unfortunately, the new spatiotemporal background is extremely complicated mathematically. There are no longer four true independent variables. Thus this interesting form of theory has seemed essentially *unworkable*. Here, too, variational principles operate in incomplete form.

In the final quantum-mechanical phase of today, there is a mystical employment of variational principles and optical analogies of classical type, in order to pass directly from the approximate dynamical interpretation of the atom (proton and electrons) to the corresponding wave equation of Schrödinger. This is a powerful engine for making calculations, and such calculations satisfy many theoretical physicists, for they lead to new power and progress. Says Dirac in this connection, "*The only object of theoretical physics is to calculate results that can be compared with experiment,* and it is quite unnecessary that any satisfying description of the whole course of the phenomena should be given" [12].

But nothing, it seems to me, is more evident than that the method proposed for making these calculations has not yet been properly placed in the wider scheme of the universe as a whole. In fact, I believe that there is a definite possibility of using electromagnetic space-time for this purpose. This possibility should be studied most carefully.

It is the principle of sufficient reason and the allied theory of groups that seem to me most significant for physical speculation. I have ventured to formulate this aspect as follows [13].

PRINCIPLE OF SUFFICIENT REASON. If there appears in any theory T a set of ambiguously determined (that is, symmetrically entering) variables, then these variables can themselves be determined only to the extent allowed by the corresponding group G. Consequently any problem concerning these variables, which has a uniquely determined solution, must itself be formulated so as to be unchanged by the operations of the group G (that is, must involve the variable symmetrically).

HEURISTIC CONJECTURE. The final form of any scientific theory T is: (1) based on a few simple postulates; and (2) contains an extensive ambiguity, associated symmetry, and underlying group G, in such wise that, if the language and laws of the theory of groups be taken for granted, the whole theory T appears as nearly self-evident in virtue of the above principle.

Indeed it may be true in some mystical sense that "God thinks multi-dimensionally [14], whereas men can only think in linear syllogistic series, and the theory of groups is the appropriate instrument of thought to remedy our deficiency in this respect." Mathematics seems to be basically occupied with model making, and the basic characteristic of a good model is found to be in its elegance in a group-theoretic sense.

But it may be asked, What is the role of variational principles, such as the principle of least action? Are they not equally important and basic? I believe that they are, in that the language of variational principles forms a kind of higher symbolism, above the calculus, so that, while the calculus is fundamental in dealing with physical theories, its processes may be replaced by more compact expression in variational terms. The true significance of variational principles in special cases remains extremely obscure. For example, an arbitrary set of ordinary differential equations is readily given variational form in a limited part of the domain of the variables. It is only the fact that there is a single explicit variational form available in the entire domain of the independent variables that is really significant. Possibly this interesting situation indicates that the basic importance of variational principles will be found to be *topological.*

This leads me to remark that topology deserves to obtain a more prominent position in physical theories than it has yet obtained. Topology may be defined as the rudimentary form of geometry, which is not concerned with shapes or size, but rather with the *connections* of geometric entities of various dimensions.

Thus in the mathematical patterns to be utilized in physics in the future, it may be conjectured that group theory, together with the allied principle of sufficient reason, variational principles, such as the principle of least action, and topology, will play an important part.

Physical theories have been, and always will be, mathematical theories; and the history of physics show inversely that those mathematical theories which have not originated in a physical model but have been directly created by the human mind, are likely to be exemplified later on in physical applications.

It will probably be the new mathematical discoveries which are suggested through physics that will always be most important, for, from the beginning, nature has led the way and established the pattern which mathematics, the language of nature, must follow.

For References see p. 312.

VI

THE MAGNETIC APPROACH TO THE ABSOLUTE ZERO OF TEMPERATURE

By PETER DEBYE

Cornell University

We measure absolute temperatures in "degrees Kelvin" (indicated by the letter K), a scale which has its zero point 273° Celsius (or Centigrade) below the freezing point of water. This point is believed to be the absolute zero of temperature inasmuch as it is the lowest limit of cold which we can approach but never attain.

The customary method for obtaining low temperatures is that of gas-liquefaction. If we want to do experiments at a temperature of about 90° K, which is equal to —183° C., we liquefy air, let it boil under atmospheric pressure, and we have a bath of the desired temperature. To prepare a bath boiling at 20° K, we liquefy hydrogen. Keeping such a bath boiling under reduced pressure will decrease the temperature further and when this is done with liquid hydrogen a temperature of 14° K is reached, at which the hydrogen solidifies, and its practical use for cooling other substances becomes very limited. We can, however, look for another gas, which is more ideal even than hydrogen, that is one consisting of molecules which attract each other with smaller forces than hydrogen molecules do. Such a gas is helium. Under atmospheric pressure liquid helium boils at 4° K. W. H. Keesom in 1932 [1] kept liquid helium boiling under a pressure of only 0.0036 mm. of mercury and obtained a temperature of 0.71° K, which is essentially the lowest that can be reached by this method.

This temperature can be considered the practical limit obtainable with liquefied gases, because, for one thing, even a substantial further decrease in pressure would result in a very insignificant additional decrease of temperature. On the other hand no gas exists which is more ideal than helium, and since nowadays we are well informed about the structure of atoms, it can be predicted that such a gas is not found in the periodic system of the chemical elements. The sole exception to this would be a gas consisting of neutrons—the neutron being the element which occupies the zero position in the periodic system, with atoms consisting of bare nuclei only, carrying no charge, and with a mass nearly equal to that of the hydrogen nucleus. Such a neutron gas can be made, but because of its peculiar atomic structure it diffuses through any container in which we try to keep it. So, obviously, the classical methods used with common gases are not suitable.

Accordingly the question arises whether another procedure can be found which will enable us to come nearer to absolute zero than the 0.7° K achieved by Keesom. This question was considered and answered independently in 1926-27 by W. F. Giauque in Berkeley (Cal.) and P. Debye, at that time in Zurich [2].

The simplest way to cool a gas, at least in theory, is to let it expand adiabatically. In this process no heat exchange is allowed, work is performed by the gas and the equivalent of this work has to be taken out of its energy content. As long as the gas can be considered ideal its energy is the kinetic energy of its molecules only. The molecules are slowed down in the process of expansion, and the gas becomes colder. Now the kind of work which the molecules are induced to deliver is of no consequence, it need not be mechanical as in this expansion experiment. We can therefore look for a system which can perform some other kind of work. However, we have to be careful in our choice; the system must be comparable to an ideal gas, which means that the change in external conditions in the course of the experiment should have no influence on its energy content if this special experiment is so conducted that the temperature of the system is kept constant. Only under these conditions can we be certain that all the work performed in the adiabatic experiment will be associated with changes in temperature and nothing else. Such an ideal system in the strict sense of the word will, of course, not exist, just as a strictly ideal gas does not exist in nature. But the nearer we can approach the ideal behavior the more efficient the cooling experiment will be.

Led by such considerations Giauque and I each proposed the use of some paramagnetic salts, which had been investigated previously as to their magnetic properties at helium-temperatures in the laboratory of Kamerlingh-Onnes in Leiden. Thus we substituted magnetic for mechanical work.

Faraday discovered that all substances can be separated into two classes with respect to magnetism. Those in the first class are repelled if brought in the neighborhood of a magnetic pole. He called them "diamagnetic," and the great majority of all known substances belong to this class. However, a small minority show the opposite behavior and are attracted to a magnetic pole as a piece of soft iron is, though for the most part much less strongly. Faraday called this second class of substances "paramagnetic" and some of these are appropriate to the cooling experiment.

In 1895 P. Curie [3] found by his experiments on oxygen gas, which is paramagnetic, that the average magnetic moment induced in such a molecule by a magnetic field depends on the temperature of the gas. The higher the temperature, the smaller the moment. Langevin [4] found the explanation for this behavior. He assumed that the oxygen molecule in its natural unperturbed state is already a permanent magnet. The effect of an applied magnetic field, then, is merely an attempt to line the axes of these elementary magnets up in the direction of the field, just as the needle of a compass is orientated in the earth's magnetic field, but this orientation is very imperfect, since it is constantly being strongly disturbed by the thermal motion of the molecules. The observed resultant effect is characterized by an equilibrium between the ordering effect of the field and the disturbing effect of the thermal motion. Langevin was able to show that the law of temperature dependence, discovered experimentally by Curie, could be derived from this picture. However, without going into detail, it is obvious that the observed average moment has, indeed, to decrease with increasing temperature.

In Curie's experiment, as well as in a large number of other experiments carried out with a variety of paramagnetic substances, the magnetic moment always proved to be proportional to the field strength. According to Langevin's picture this means that, in all these cases, the thermal disturbance is preponderant, and as a result we have only an incipient orientation. However, if experiments were conducted at low enough temperatures, a saturation effect might be expected, since the utmost a magnetic field can do is to arrange all the elementary permanent magnets parallel to the field. This was the effect detected in the Leiden laboratory with some paramagnetic salts, as, for instance, gadolinium sulfate [5]. We have here, of course, an important confirmation of Langevin's fundamental idea, but this does not necessarily imply that all the details of his picture must likewise be accepted. As a matter of fact it is now known that the elementary magnet, in such cases as we have to consider here, is the electron itself, which has its spin and, connected with this spin, a permanent magnetic moment. We know also that very special and peculiar rules hold for the possible orientations of this moment. In one electron the orientation can be only parallel or antiparallel to the direction of the field; there are no intermediate positions. At this stage of our argument, however, it is sufficient to realize, first, that at temperatures obtainable with liquid helium and in some appropriate substances a high percentage of paramagnetic saturation can be accomplished with magnetic fields of 20,000 to 30,000 Gauss, this last field being roughly the practical limit of the field strength for such electromagnets as can be made available in the laboratory. Due to this circumstance a significant amount of work can be performed in magnetizing the paramagnetic substance, this work being done in overcoming the opposing electromotive force in the coil surrounding the iron-core of the electromagnet during the time the electric current is brought to its final value. Second, we realize that, even at helium temperatures, paramagnetic substances exist which are still very nearly ideal, due to the fact that the interaction of the elementary magnets with their surroundings is very small (Fig. 49). As a result the energy content does not depend on the magnetization, but only on the temperature, in the same way and for the same reason that the energy of an ideal gas does not depend on its density.

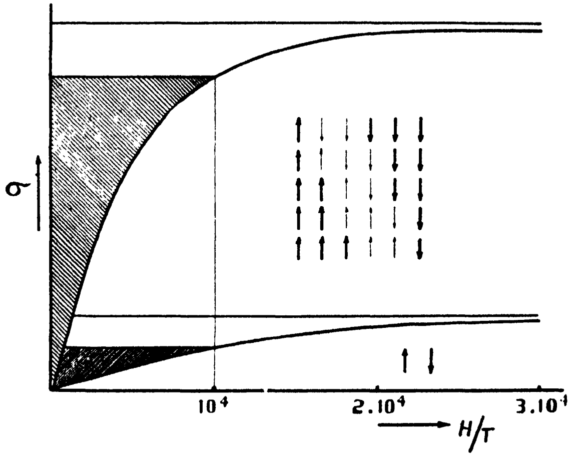


FIG. 49. The magnetic moment σ of the sample is represented as a function of the quotient magnetic field strength H divided by the temperature T. This special relation holds for *ideal* magnetic substances. For large enough values of H/T the magnetic moment tends to a limit (saturation). One of the 2 curves refers to atoms with 1 free magnetic electron (trivalent Ti) the other to atoms with 5 free magnetic electrons (trivalent Fe). Possible arrangements of the spin-axes are indicated by the arrows. The shaded surfaces in the diagram represent the amount of work done in magnetizing up to a value of 10,000 for the quotient H/T.

With these considerations in mind we are now prepared to describe the procedure of magnetic cooling.

A suitable paramagnetic substance is precooled with liquid helium, under an appropriately low pressure, to a temperature of roughly 1° K. The substance used has no immediate contact with the liquid helium, but is separated from it by an interspace filled with some helium gas which maintains the heat contact. The whole arrangement is situated between the poles of an electromagnet (Fig. 50).

As a first step the electric current is established in the coils of the electromagnet, bringing the magnetic field between its poles, to say, 25,000 Gauss. During this magnetization process, work is done in the paramagnetic substance (comparable to the work associated with the compression of a gas) and, since the substance is ideal and the temperature is kept constant by the surrounding helium bath at 1° K, an equivalent amount of heat is transported to the outside and spent in evaporating some liquid helium.

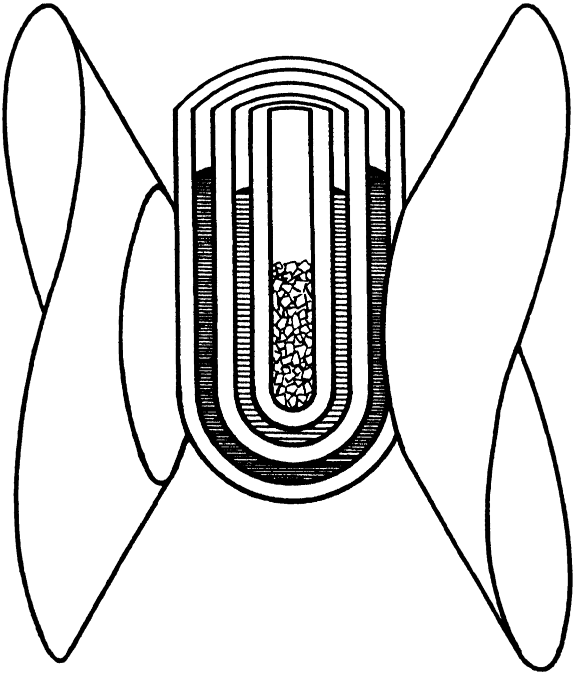


FIG. 50. In this schematic picture the paramagnetic substance in the form of a crystalline salt is represented between the poles of an electromagnet. It is surrounded by an inner sheet of liquid helium, which is protected by an outer sheet of liquid hydrogen.

As a second step the helium gas, which was necessary for the heat transport, is removed, leaving the paramagnetic sample isolated from its surroundings at a temperature of 1° K and under the stress of the established magnetic field.

The third and last step is comparable to an adiabatic expansion of a gas. Either the instrumental arrangement containing the sample is removed from the field or the field is eliminated by decreasing the magnetizing current to zero. During this process the substance has to deliver work and since no heat flow is allowed, it has to decrease its temperature below the initial temperature of 1° K with which we started.

At the time when a trial of the magnetic method was suggested nothing definite could be predicted about the final temperature at which the substance would settle at the end of the procedure. From the fact that the energy of the vibrations of the atoms in the lattice of the crystal is very small at 1° K, much smaller than the magnetic work involved in the magnetizing process, it seemed probable that good results could be obtained. But it was also evident that, since the atomic vibrations are so very unimportant, all would depend on the magnetic properties of the substance below 1° K, and these in turn could only be adequately investigated by performing the experiment which was the only means of getting substantially below this temperature.

For a long time I did not succeed in securing the funds necessary for the rather elaborate low-temperature and magnetic equipment involved. Giauque was more successful in Berkeley. He has patiently built up a laboratory well adapted to his purpose. Three men—Giauque; F. Simon, working originally in Breslau, later in Oxford, and supplementing his low-temperature apparatus by the use of the big electromagnet of Cotton at Bellevue near Paris; and W. J. de Haas, one of the two successors of Kamerlingh-Onnes in the splendidly equipped Leiden Laboratory—are mainly responsible for the advance made in the years between 1927 and the beginning of the war.

At first progress was very slow. This was partly due to the fact that the paramagnetic substances originally used in the experiment turned out to exhibit deviations from their ideal behavior only a short distance below 1° K. This was also true of gadolinium sulfate, thoroughly investigated by Giauque [6], and of other rare earth salts—for instance, of cerium fluoride on which W. J. de Haas and E. C. Wiersma [7] report, as the result of an experiment performed June 15, 1933, a final temperature of 0.13° K in a process starting with 27,600 Gauss at 1.35° K and ending with 850 Gauss.

A very important improvement was made when de Haas began using elements of the iron-group in the periodic system occurring in trivalent ionic forms in alums. Again, with Wiersma [8], he began work with the alums in December, 1933, and on March 7, 1935, an experiment performed on potassium-chromium-sulfate (diluted with another non-magnetic alum) yielded a final temperature of 0.0044° K. The volume of the sample was 56 c.c., and the demagnetization process was started in a field of 24,075 Gauss at 1.174° K and ended at 1 Gauss.

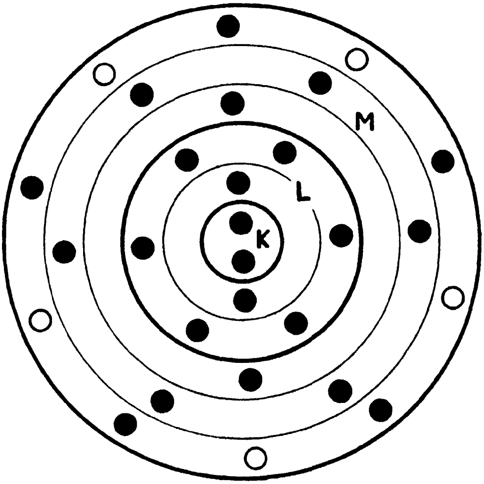


FIG. 51. A schematic representation of the distribution of the 23 electrons in a trivalent Fe-ion. The K-shell is completed by 2 electrons; the two compartments of the L-shell are filled with 2 + 6 electrons; the three compartments of the M-shell are able to receive 2 + 6 + 10 electrons. Only 2 + 6 + 5 are available and, therefore, 5 places indicated by open circles are left unoccupied.

As previously mentioned, the final temperature at the end of the process is determined by the existing deviations from an ideal paramagnetic state, due, in their turn, to interaction of the magnetic atom with its surroundings in the crystal. But before attempting to visualize those interactions it becomes necessary to discuss the processes involved from a more fundamental standpoint than before. This can be done by centering the argument around the entropy and its interpretation, as a measure of the disorder, according to L. Boltzmann.

Let us consider the special case of ammonium-iron-alum (NH_4 Fe- $(SO_4)_2 + 12\ H_2O$) in which the trivalent iron-ion carries the paramagnetic properties. In its trivalent state the nucleus of this atom is surrounded by a cloud of 23 electrons, 2 in the K-shell, 2 + 6 in the two compartments of the L-shell, and 2 + 6 + 5 in the three compartments of the M-shell (Fig. 51). In the third compartment of this last shell a maximum of 10 electrons could be accommodated, and hence it happens that the last 5 electrons can be considered free with respect to the orientation of their spin-axes in a magnetic field. It is easily seen that since only two positions, parallel or antiparallel, are allowed for each electron, six different combinations can occur, which lead to magnetic moments that are 5, 3, 1, —1, —3, or —5 times the elementary electronic moment (compare Fig. 49). In a very small magnetic field the energy of each of these combinations with respect to the magnetic field can be considered as being the same. The Fe-ions can make a free choice between the six possibilities and, therefore, will be in a state of disorder, which we can characterize by assigning to it the number 6. Boltzmann's fundamental idea now is that the number characterizing the disorder of a system has an intimate relation to its entropy. He shows that if we accept the existence of such a relation the entropy as defined in the usual way (by the thermodynamic relation that for a small change its differential is equal to the heat imparted to the system divided by the absolute temperature) must be proportional to the natural logarithm of that number. Considering one atomic weight of Fe-ions (56 gr.) the proportionality factor is the gas-constant R with a value of roughly 2 calories.

We conclude in this way that the entropy content of an atomic weight of free Fe-ions connected with its paramagnetic properties will be $R \ln 6$ (Fig. 52). This is at the same time a measure of the natural disorder in those atoms.

As soon as the magnetic field is increased in strength the six combinations will become differentiated from each other by different energy values, since turning a magnet in a magnetic field from a parallel to an antiparallel position requires work. Consequently our Fe-ions will now show a preference for the positions with lowest energy. In very strong fields the state having +5 times the elementary amount (moment and field parallel to each other) will occur predominantly. But this being so the disorder number will now have become very nearly equal to 1 and the entropy, being proportional to the logarithm of this number, very nearly 0.

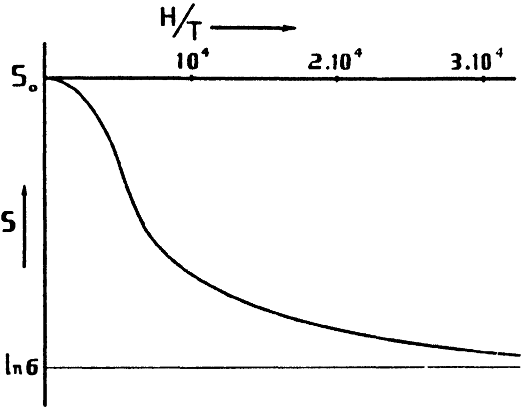


FIG. 52. The entropy S is represented as a function of H/T for an atomic weight of trivalent Fe-ions, considered to behave ideally. At constant temperature and with increasing field strength the entropy decreases from a value S_0 to a final value $S_0 - R \ln 6$.

Thus it is seen that in applying a magnetic field to a paramagnetic substance we decrease the entropy, and if the process is carried out at constant temperature this obviously means that heat is generated in our system and has to be carried to the outside. However, we can say with equal right that the application of a magnetic field at constant temperature decreases the disorder or increases the order in the system (Fig. 52).

In our third and last step of the magnetic cooling process the field was removed adiabatically. In such a process no heat transfer exists, therefore the entropy and with it the disorder remain constant. Our main question as to the final location of our sample on the temperature scale can now be answered by saying that we will find it in a state having the same diminished amount of natural disorder as we were able to enforce with the help of the magnetic field at a higher starting temperature.

How and why the order in a salt like our iron-alum still increases significantly if we come nearer to absolute zero beyond the point where the disorder due to the atomic vibrations has already been eliminated was discussed by J. H. van Vleck, M. H. Hebb, and E. M. Purcell [9]. In a general way the effect is due to interactions of the paramagnetic atoms with their surroundings in the crystal, and it is analogous to an effect of the same kind in gases, their transition to the liquid state being connected with the existence of the intermolecular forces.

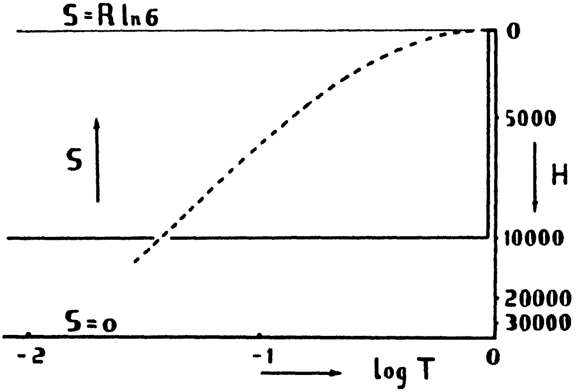


FIG. 53. The entropy of an atomic weight of trivalent Fe-ions is represented as a function of the temperature plotted on a logarithmic scale. The arbitrary constant in the entropy has been adjusted so as to make S = o in an infinitely strong field. The first step in the process performed at constant temperature T = 1 is represented by a vertical line going from a point characterized by H = o, to a point corresponding to H = 10,000, the field supposedly applied. The last step, representing the adiabatic demagnetization which involves constancy of the entropy, is indicated by a horizontal straight line. Another dotted line shows how we expect the entropy to change with temperature in an unmagnetized sample as a result of atomic interaction. The meeting point of these two lines determines the final temperature.

X-ray evidence has shown that in our alums every one of the two positive ions is highly hydrated [10]. Six water molecules occupy the corners of an octahedron surrounding the iron ion and, therefore, put its electronic cloud under an electrical strain of special symmetry related to the positions of the water molecules (Fig. 54). The result is that two distinct electronic distributions with different energy content become possible. Expressed in the language familiar to the chemist this means that, under the special conditions prevailing in the crystal, the trivalent iron ion can occur in two different modifications. These two modifications differ not only in energy content but also in their magnetic reaction to a magnetic field (Fig. 55). If such a field is applied, one of them, which we will assume to have the higher energy content, can separate into four, the other into two different states. The energy difference of the two modifications is small, even in comparison with the average kinetic energy of a free atom at 1° K. At this temperature we shall therefore find, on the average, ⅔ of the Fe-ions in the form of the higher and ⅓ in the form of the lower modification. As soon, however, as in lowering the temperature we pass through the region in which the average energy of a free atom becomes

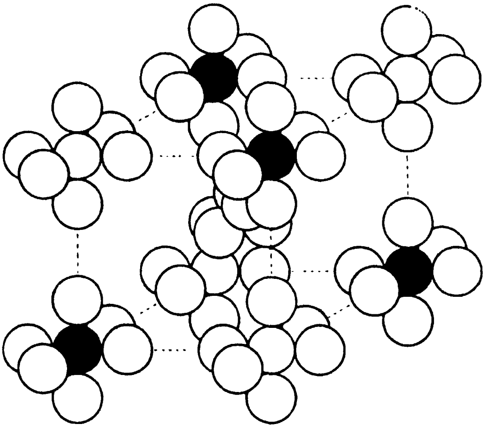


FIG. 54. One elementary cell of Fe(NH$_4$) (SO$_4$)$_2$ + 12 H$_2$O is represented. The center of the cell is occupied by the SO$_4$-group, the O-atoms being at the corners of a tetrahedron. In the picture they hide the central S-atoms from our view. Four of the eight corners are occupied by Fe-atoms, the four others by NH$_4$-groups. The Fe-ions are represented by black balls, the NH$_4$-groups by white balls. Every one of these ions is surrounded by six H$_2$O-molecules, also represented each by a white ball.

equal to the energy difference of the two modifications the Fe-ions begin to segregate more and more in the state of the lower-energy modification. Below this region in the temperature scale we shall soon find that nearly all the Fe-ions belong to the lower modification, which can be split in only two different states by the application of a magnetic field. In this way the disorder decreases naturally, without the help of a magnetic field, from R ln 6 to R ln 2 (Fig. 56).

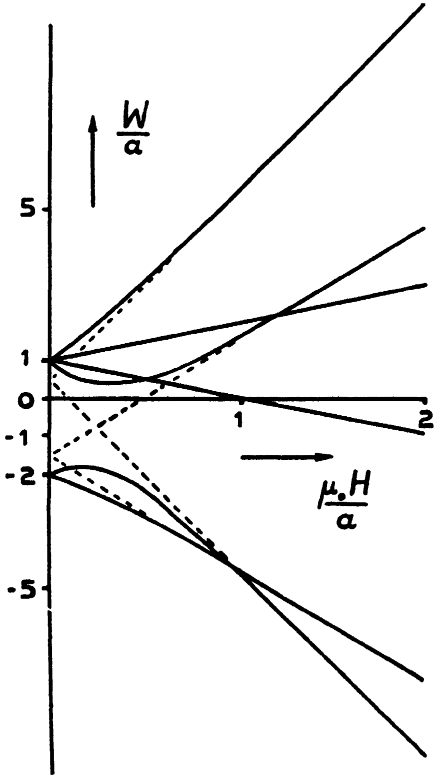


Fig. 55. The energy W of the 6 different states of a trivalent iron ion exposed to the strain exerted by the surrounding H_2O-molecules has been represented as a function of the strength H of the magnetic field. The quantity a is an energy constant characterizing the strain and μ_0 is the magnetic moment of the spinning electron. For a finite field 6 states with different energy content exist. In the limit for H-o the 6 states collapse into 2, one of these composed of 4 and the other of 2 representing the 2 modifications mentioned in the text.

As a consequence of this picture we must expect that in the transition region an anomaly in the specific heat will occur since in passing through this region from lower to higher temperatures energy will have to be provided to the atoms transferring from the lower to the higher modification. The salt has been investigated by F. Simon and his associates N. Kürti, P. Lainé, B. V. Rollin [11], as well as by W. J. de Haas and associates, H. B. C. Casimir and D. de Klerk [12]. The anomaly is very prominent around a temperature of 0.07° K (Fig. 56).

This discussion still leaves us with a salt which tends to a final disorder value of R ln 2 for temperatures far enough below 0.07° K. On the other hand it is possible to decrease the disorder more than the difference of R ln 6 and R ln 2 by applying a strong enough magnetic field at about 1° K.

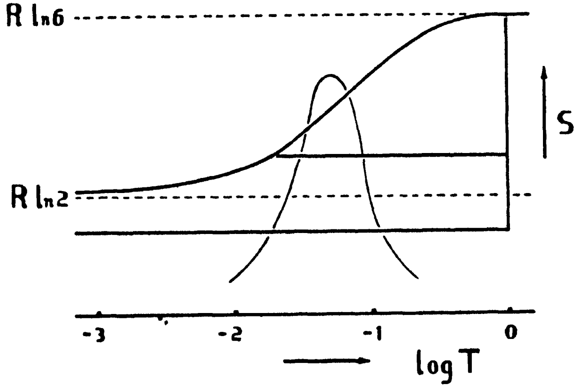


FIG. 56. The figure indicates how the entropy of the iron-ammonium-alum is expected to decrease from R ln 6 to R ln 2 if no other influence than that of the H_2O molecules is taken into account. It illustrates the theoretical difficulty for fields strong enough to decrease the entropy more than R ln 6 — R ln 2. The bell-shaped curve represents the anomaly in the specific heat of the unmagnetized sample connected with the S-shaped entropy curve.

Should no mechanism exist by which the lower modification could get rid of its disorder in a natural way at still lower temperatures, we would be able to perform a process ending at the absolute zero of temperature itself. We are convinced that a process of this kind is impossible and, as a matter of fact, another interaction not yet mentioned takes care of our difficulty. All the Fe-ions in the crystal are coupled by magnetic interactions and although these interactions are very weak they become important at temperatures as low as 0.01° K to 0.001° K. They have been considered by Waller (13) and consequently by many other investigators. The effect of the interactions is that not the single Fe-ions but the whole crystal, comprising all the ions together with their interactions, has to be treated as one system. To use again the language of the chemist, it will be found that the crystal can exist in many modifications, which, because of the magnetic interactions involved, will have different energy contents. It is evident that under these circumstances the whole crystal in approaching absolute zero can and will decrease its natural disorder by falling back more and more on the modifications of lowest energy content. In this way we realize that, ultimately, the practical low temperature limit of the magnetic process is determined by the magnetic interactions of its constituent paramagnetic atoms. This interaction can be decreased by placing those atoms farther apart, a method employed by de Haas when he diluted magnetic with non-magnetic alums. But in this respect we cannot go very far, and a simple calculation, comparing the magnetic energy of interaction with the thermal energy of a free atom, leads us to believe that the limit is situated at a temperature of the order of 0.001° K. As shown by de Haas, this checks with the practical achievement as it stands today.

From a theoretical point of view not enough is known about the magnetic structure of the different crystalline modifications. F. Simon reports that experimentally ammonium-iron-alum shows a Curie-point at 0.34° K at which the crystal becomes ferromagnetic, that is, begins to behave much like an ordinary piece of iron at room temperature.

In the preceding discussion I have considered temperatures ranging from a few thousandths to one degree absolute, but have given no indication how such low temperatures can be measured. In the temperature range accessible with the help of liquefied gases such measurements are based on readings of the gas thermometer. However, in the new range below 1° K such an instrument is useless. This is readily seen in computing the vapor pressure of liquid helium, for which a formula exists containing constants derived by Keesom from his experiments. At 0.7° K the calculated vapor pressure in mm. of mercury is $2.93.10^{-3}$; however, at 0.2° K the formula yields $3.58.10^{-14}$ mm. and at 0.1° K, $5.88.10^{-31}$ mm. This means that at 0.2° K we may still expect to find nearly 2 million atoms per cc., but at 0.1° K this will be reduced to $5.6.10^{-11}$. In other words, every atom of the gas may be expected to occupy the space of a 26-meter cube. It is obvious that the gas thermometer under these circumstances is useless indeed. It is also clear that, at the low temperatures we are using, no gas will remain in the interspace of the Dewar containers used in the experiments. Everything will be condensed on the walls, and the containers will automatically become perfect insulators so far as heat conduction by the gas in the interspace is concerned.

The method which can be used for temperature measurement is based on measurements of the magnetic properties of the samples. The lower the temperature the stronger is the magnetic moment which can be induced in a sample by a small magnetic field and the stronger are the forces with which such a sample is attracted into those parts of an inhomogeneous field where the lines of force are crowded. By measuring such forces we determine the so-called susceptibility of the sample, that is, the moment per gram induced by a field of unit strength. In ideal magnetic substances Curie's law holds, which states that the susceptibility is proportional to the reciprocal of the absolute temperature. So if we had an ideal substance and had found at the end of our process a susceptibility say 100 times larger than that measured at 1° K, we would know that a temperature of 0.01° K had been reached. However, the substances used are not ideal and we must discover how to measure the absolute temperature in spite of that. The same question rises in connection with the gas thermometer, since after all no gas is strictly ideal. It has been answered in this case and W. H. Keesom [14] has indicated how the reasoning can be adapted to the establishment of a magnetic temperature scale.

Suppose we want to know the average absolute temperature that can be achieved, starting with a field between 10,000 and 11,000 Gauss at a temperature of 1° K and demagnetizing the sample completely. It is easy to find out how much decrease in disorder (entropy) is connected with the increase in field strength from 10,000 to 11,000. This is equal to the heat produced in this step divided by the temperature (1° K), or it can be derived from the magnetization curve. Since the processes which end in the two final states, beginning with 10,000 or 11,000 Gauss, are adiabatic, we know that these two states will have the same difference in disorder. But this means that if we are able to measure how much heat we have to apply in order to change from one final state to the other, our problem is solved, since the disorder (entropy) difference is equal to this amount of heat divided by the temperature we wish to determine (Fig. 57). From an experimental point of view the question is how the sample can be heated in its low-temperature state in a controllable way and how the two final states can be identified.

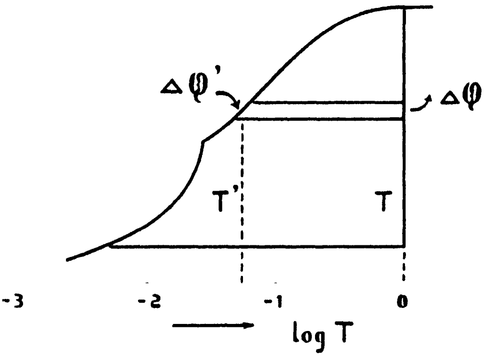


Fig. 57. The diagram represents the entropy of unmagnetized iron-ammonium-alum as a function of the temperature. Increasing the field from 10,000 to 11,000 Gauss at a constant temperature T = 1, releases the heat ΔQ. To heat the unmagnetized sample over the same entropy interval necessitates the introduction of the heat ΔQ′. According to the thermo-dynamic definition of entropy we have

$$\Delta Q'/T' = \Delta Q/T$$

from which T′ can be calculated.

The last part of this question can be answered at once; we have only to measure the susceptibility or, more practically, the force with which the sample is drawn into a small inhomogeneous field. Heating the sample can be accomplished by exposing it to an alternating magnetic field, an experiment which is interesting for its own sake, since it teaches us about the relaxation time involved at our low temperatures in the process of inducing a magnetic moment by a field. Very interesting experiments in this direction have been initiated by C. J. Gorter [15]. Another ingenious way of heating was originally applied by F. Simon [16]. He heats with γ-rays, which pass through the walls of the Dewar containers and heat the sample by the fraction which it absorbs. In this method of heating the amount absorbed is proportional to the time required by the sample to pass from one final state to the next as characterized by their susceptibilities. In describing this method of temperature determination we made use of an interval of 1000 Gauss, but this, of course, was only for the sake of argument, and it is obvious that the procedure can be adapted to yield definite temperature values for definite final states.

Looking back on the extension of the accessible temperature range to be accomplished with the help of the magnetic method, we see that it amounts to a factor of about 1000 in the temperature scale (1° K to 0.001° K). In connection with a remark made long ago by Lord Kelvin, the quotient of two absolute temperatures rather than their difference has here been introduced to evaluate the achievement. Implicitly this means, in agreement with Lord Kelvin, that the logarithm of the temperature is considered a better scale than the temperature itself. Low-temperature work in all its aspects bears out the supremacy of the logarithmic scale, which incidentally has the advantage of removing the absolute zero to $-\infty$, certainly a help to every student.

Finally, we can speculate about the chance of extending the experimentally accessible range by another factor 1000 toward the absolute zero. Characteristic for the range covered by the method of gas liquefaction is the fact that we are dealing with the atoms of molecules and their motion. We can say that at helium temperatures we have already succeeded in removing nearly all the disorder connected with such motion. In the next stage covered by the magnetic method the disorder connected with the electrons and their spin in the interior of the atom is essentially removed, in the electronic cloud, surrounding the nucleus. Speculations have been made about the possibility of removing a subsequent kind of disorder which may exist in the nucleus itself and may be connected with the nuclear magnetism. Encouragement is derived from the fact that the magnetic units in the nucleus having magnetic moments roughly 1000 times smaller than the electronic Bohr magneton might be accessible to experiment now that the temperature has also been lowered by the same factor. However, no experiments have yet been made. The chances are that not until the war is over can further attention be paid to this aspect of the subject.

For References see p. 313.

VII

THE DRIFT TOWARD EQUILIBRIUM

By HENRY EYRING

Princeton University

INTRODUCTION

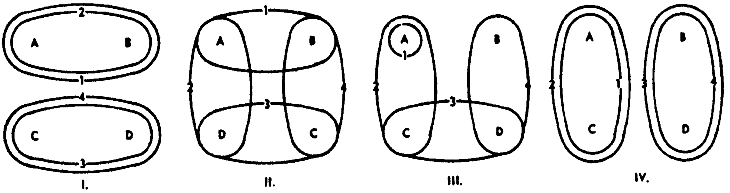
Much of the world's work is done by people who have learned their methods by watching someone else carry through a simple or complicated set of manual operations. Established procedure is modified from time to time as improved results are observed to accompany some variation. Theoretical chemistry, on the other hand, interprets experimental data in terms of atoms and molecules. Among my early memories of cattle ranching and farming I fail to recall a case where anyone was instructed as to procedure in terms of atoms or molecules. A later brief career as a practical copper miner likewise added little to a very limited knowledge of the atomic world. One thing soon became abundantly clear, however, namely, that not only apples but large blocks of porphyry are attracted to the earth—often falling with great speed and bonecrushing finality. These observations might have led even a less impressionable person to ponder the advantages and disadvantages of bombardment by individual atoms as compared with atoms in bulk.

There are many possible compromises between the study of matter in bulk and the focussing of attention on atom mechanics. Many scientifically trained college graduates have been observed to live reasonably unfrustrated lives after having forgotten about valence. This is possible because the atoms inside us—unlike human beings—do not feel obliged to retaliate for neglect by showing a reciprocal obliviousness to their appointed task. Each carbon atom untiringly directs four half bonds—each consisting of an electron—toward four neighbor atoms each of which reciprocates by extending a similar electron half bond to overlap that of the carbon. This pairing of electrons in bonds was pointed out and clarified by G. N. Lewis long ago. Quantum mechanics amplifies the picture by showing that only if the two electrons are spinning on their axes in opposite directions will they consent to form the two halves of an electron pair bond. A third electron attempting to enter such a bond is *persona non grata* and is obliged to seek a second empty or half-empty bond to move into. As often as not this available bond, or orbital, is higher up in the atomic hotel farther removed from the cheering warmth of the positive charge at the atom's center. Most of the time this electron exclusiveness works no particular hardship. Each atom simply fills up its orbitals with electrons until they balance the central positive charge. The electrons go into the lower orbitals two at a time, since the alternative is to move upstairs. In the topmost partially occupied floor, however, they spread out one to a room if this is possible without "promotion," i.e., without moving upstairs. This is because even electrons with opposite spins are repellent to each other, due to their like charges, and only suffer each other's company to escape an even less desirable situation.

Now an atom's covalence is just the number of these half-filled orbitals each of which can pair with the half-filled orbital from some neighbor atom. This inter-atom pairing occurs as a result of the fact that each electron thereby replaces the attraction of one nucleus by that of two. If one atom attracts the electron pair much more than the other the resultant bond, with the electrons shifting to one side, forms an electrovalence. Besides these covalences and electrovalences, an atom having a completely unoccupied orbital at the level of the covalences may share it with some atom having two electrons in an outer orbital to form a Werner coördination compound.

Any collection of atoms will thus organize itself into molecular arrangements with an appropriate assortment of covalences, electrovalences, and coördination bonds.

The most stable molecular arrangement shifts with the temperature, pressure, and composition of the system. And it is these shifts toward equilibrium that are retarded by electron exclusiveness. If two covalently bound atom pairs decide to change partners they are up against a difficult proposition. The successive stages in the reaction AB + CD → AC + BD are represented in Figure 58:



I. Initial state. II. and III. Molecular structures contributing to activated state. IV. Final state.

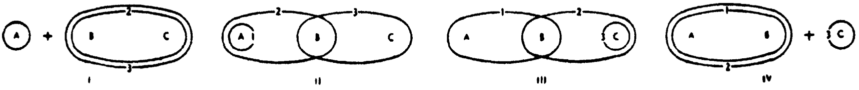
Fig. 58. Illustrating contributing states (I, II, III, IV) in the reaction AB + CD → AC + BD

In Figure 58 the letters represent each of the four atoms minus its bonding electron. The bonding orbitals are represented by the ellipses (or circles) labeled by arabic numerals. Equivalent orbitals are drawn slightly different in size to avoid overlapping. The numbers are placed where the electrons tend to linger, that is, in between two atoms where they can be near both positive nuclei. The spin angular momenta of electrons in orbitals 1 and 3 are in the same sense—say clockwise—throughout the reaction and opposite to the spins in orbits 2 and 4.

The activated configuration II indicates a state half way between the initial and final state and in quantum mechanics it is treated as a "superposition" in equal amounts of the initial and final states. Electronic configuration III and the three like it, not shown here, in which the circular orbit shifts in turn to B, C, or D, represent the behavior of the electrons in the activated state during a considerably smaller fraction of the total time than does II. These configurations of type III are polar and are frequently neglected in ordinary calculations of the energy of the activated state.

The activated configuration II—a mixture of the initial and final configurations I and IV—is said to resonate between the initial and final configuration. Those who are familiar with the stabilizing influence of resonance in molecules, such as benzene, may be surprised at the fact that the unstable activated state shows greater resonance than the stable initial and final states. This situation arises because the resonance in the activated state, which, just as in ordinary molecules increases stability, is accompanied by an increased repulsion between the positive nuclear charges which overbalances resonance stabilization by about one quarter the strength of the bonds broken by the reaction as it proceeds in the exothermal direction.

The three atom reaction A + BC → AB + C is represented diagrammatically in Figure 59:



I. Initial configuration. II. and III. Activated configurations. IV. Final configuration.

Fig. 59. Illustrating contributing configurations (I, II, III, IV) in the atom reaction A + BC → AB + C.

Actually all the configurations I to IV enter into the activated configurations. Because the activated configuration is linear, nuclear repulsions are not large and the activation energy is a few large calories only.

Fight against it as he may, the theoretical chemist describing molecules is cast in the role of Lewis Carroll. If with Alice we shrank a hundred millionfold and increased our heartbeat ten million million times, we would step into a world which would make the Mesozoic with its dinosaurs and other horrors seem like a quiet Sunday afternoon. We would see long strings of atoms held together by electron pairs jumping back and forth between neighbors. We would note that atoms with a valence of one always terminate chains; bivalent atoms are always inside chains; tri- and quadrivalent atoms always cause chains to branch; and the branches frequently end after adding a single monovalent atom. In fact, it is probable that if the plasmodium parasite, responsible for malaria, ever saw a quinine molecule approaching it would "give up the ghost" in sheer terror. The most startling thing would be to see the molecules floating through space turning and twisting and banging into other molecules like living things. In the typical molecular collision very little happens except a moderate change in direction and velocity of the colliding molecules. The average velocity of molecules in the gas phase is the velocity of sound. The molecules are, of course, the messengers which carry the slightly increased impulse that is the sound wave.

While the molecules survive the average collision without any obvious damage, about every hundred-million-millionth molecule at ordinary temperatures comes whizzing by as though shot out of a gun, and when it hits another molecule something happens. Each of the colliding molecules may be seen to break into two parts, each of which joins with a part of the other molecule to form a new one. These new molecules depart at about half the old speed but trembling in every atom from the experience. We have witnessed a bimolecular reaction, no less. Sometimes the result of such a collision is to transfer substantially all the great relative molecular motion into the interior of a single molecule, and then we witness a convulsive writhing and twisting of the doomed molecule which is truly terrifying. Finally the molecule breaks in two and everything is relatively quiet, and we have seen a unimolecular reaction. The two pieces into which the molecule separated may contain unpaired electrons and so one or both of them after gaining sufficient speed may bump into a passing molecule and bite off a number of atoms. The mutilated remainder may in its turn continue the process. This is a chain reaction.

If we walk around among the molecules for a while and keep our eyes open we soon notice that collisions violent enough to be the undoing of certain types are passed off with only a shudder by others. Also a pair of neighbors in one molecule will be observed to separate relatively easily if they bump into the right atom or atom pair in another molecule. In fact organic chemists keeping careful records of their experience know that if some new type of molecule is split in two by certain added molecules at a certain temperature and pressure it is because the new molecule contains a certain pair of atoms as neighbors. By repeatedly testing with appropriate reagents they finally are able to establish how all the atoms are tied together and so write down the structural formula of the new molecules. In general of course, each breaking of bonds in analysis means the making of others—which is a synthesis. Thus in any particular reaction one says he is analyzing or synthesizing according to whether he is interested in how reactants break down or in the products formed.

Only a statistically minded observer would notice that the average value of one half the mass of molecules times the velocity squared is the same for all molecules, large and small, and that this energy is proportional to the absolute temperature. However, once he noticed this, and also that molecules colliding with more than a certain critical energy will react, the next steps to take should be obvious enough. They are these: (a) What is this critical energy of collision for different molecules depending on their composition and structure and (b) if molecules have a certain mean energy, i.e., temperature, what fraction of them have more than the critical energy and so will react at the next collision? In the discussion that follows we shall set forth, in so far as is possible, the answers to these fundamental questions.

As we have said there are many interesting aspects of chemistry which involve the manipulation of material in bulk to produce new products of first-class everyday importance. Under such conditions one rarely needs to consider time intervals much shorter than seconds, or distances or amounts of matter less than one can see. This macroscopic world is all one needs to know if one is willing to leave unanswered the question of how our physical world works. The fact that ultimately all chemical events involve material units that can be packed a hundred million to the inch and reorganizations which are completed in a trillionth of a second gives the entirely new scale of time and space into which our explanations must be fitted. It is our purpose now to examine, as best we can, the laws of behavior in this sub-microscopic world.

A thorough understanding of the equilibrium states of matter enables us to state the general direction a non-equilibrium system will take in shifting toward equilibrium. Accordingly, thinking in thermodynamic terms, Clausius was able to say, "the energy of the world is a constant. The entropy tends toward a maximum." However, equilibrium theory leaves entirely out of account the question of the rapidity of such changes. Thus, by being sufficiently careful, one can keep highly explosive mixtures, such as hydrogen and oxygen or hydrogen and fluorine, for indefinitely long periods. In such extreme cases, of course, the penalty for error may be serious and immediate.

Thermodynamic theory plus a knowledge of the forces between atoms enables us accurately to predict what the final equilibrium pressures will be for the hydrogen-oxygen mixture in spite of the fact that temporary changes in temperature, or in pressure, or traces of catalyst, such as platinum black, profoundly affect the time required for equilibrium to be reached. And so for each system displaced from equilibrium there is an endless variety of ways of approaching the final state, so that there is no general answer to the question how soon equilibrium will be attained. However, if the conditions are sufficiently well specified, the rate of a reaction is reproducible and predictable. Such predictions may rest on previous experience but also—in theory at least—the rate may be calculated *a priori.*

The steps involved in such a prediction for a sufficiently specified reaction are several. They include (1) a calculation of the probability of energy states in a given molecular system; (2) the interpretation of reaction equilibrium in terms of a statistics which defines exactly the component energies of the molecules constituting such a system under the specified conditions, and explicitly in terms of the configurations or atomic distances which the constituents may assume either as reacting materials or reaction products; and finally (3) a new theory of rate processes which defines the velocity of reaction in terms of the energetically easiest reaction path by which the reactants may pass from their initial atomic configurations to their final configurations as reaction products. This new theory requires the fundamental assumption that the statistical theory of equilibrium involved in (2) is applicable to systems of reactants or products and those energy-rich configurations which intervene in the process of reaction.

We shall consider each of these aspects of the problem in turn and then proceed to an examination of some examples illustrative of the whole procedure. It will emerge that this method of procedure is applicable to simple and also to complex chemical reactions. The method is equally applicable to drifts toward equilibrium that are involved in many physical processes promoted by an external force, such as are involved in viscous and plastic flow, diffusion, and electrical conductivity. More recently the same method of procedure has been applied to a variety of biological processes, notably the emission of light by luminescent animals as affected by both temperature and pressure and by inhibitors, such as poisons and narcotics. The powerful nature of a theoretical advance based upon statistical theory can thereby be revealed in chemistry, physics, and biological science.

THE PROBABILITY FUNCTION FOR THE DISTRIBUTION OF ENERGY IN A MOLECULAR SYSTEM

An interesting way of considering energy distribution is presented below and in more precise form elsewhere [11]. The fundamental postulate of statistical mechanics can be stated as follows: *In any system of molecules (or atoms) any exactly specified way of distributing the total energy E among these molecules is just as likely as any other exactly specified way of distributing this same total energy among these molecules.* Now if from quantum mechanics or spectroscopy we know all the allowed energy levels for a set of molecules we can, if we choose, enumerate all the possible ways of distributing a fixed amount of energy E among them. A definite fraction of all these ways allocates the energy ϵ_1 in an exactly specified way in some particular part of the larger system. This fraction is then the probability that if the whole system has the energy E this particular part has the energy ϵ_i allocated in this special way. But fixing the total energy E for a system of molecules fixes the average energy per molecule and hence the absolute temperature T. This is because average energy is specific heat times temperature. Thus we have a a straightforward method of calculating the chance at any temperature that a particular part of a system will possess the energy ϵ_i in some specified way. This chance, p_i, of being in the state with energy, ϵ_i is readily found to be $p_i = a\, e^{-\epsilon_i/kT}$ (1a). Here the Boltzmann constant, k, is the high temperature specific heat of a single harmonic oscillator.

The sum of the probabilities of all states a system can occupy must add up to unity, that is, $1 = \sum_i a\, e^{-\epsilon_i/kT}$. Dividing this into the above equation one obtains for the probability of the ith state

$$p_i = \frac{e^{-\epsilon_i/kT}}{\sum_i e^{-\epsilon_i/kT}} \qquad \text{(1b)},$$

which depends only on the energy levels and the absolute temperature.

The "particular part" of the larger system having the energy ϵ_i may be any part, simple or complex, about which one can properly say it possesses this energy in a specified way. The simplest motions of systems with which definite energies can be associated are spoken of as degrees of freedom. Thus an elementary particle can move in three independent directions and so it possesses three degrees of freedom of motion, and a molecule composed of n particles possesses 3n degrees of freedom. The energy of a complicated system is, of course, the sum of the energies of its constituent degrees of freedom.

In outlining the derivation of the Boltzmann expression for p_i it was not necessary to specify anything more about the system than the allowed energy levels. Since all of equilibrium theory and almost all of rate theory can be made to follow from this expression (1) we see the essential unity and simplicity of thermodynamics and rate theory.

EQUILIBRIUM THEORY

The same set of atoms frequently can exist in more than one equilibrium arrangement. Thus, for example, a system of two hydrogen atoms and two iodine atoms besides existing as atoms can form a hydrogen molecule and an iodine molecule, or, alternatively, two hydrogen iodide molecules. The equilibrium constant, which is the chance of being in the final state over the chance of being in the initial state, is obtained by dividing the sum of probabilities for each possible energy for the final state by the corresponding sum for the initial state. Accordingly the equilibrium constant in view of (1) after canceling out the letter a is

$$K = \frac{\sum_{f} e^{-\epsilon_f/kT}}{\sum_{i} e^{-\epsilon_i/kT}} \tag{2}$$

where the upper summation is over energy levels of final state and is called the partition function for the final state, whereas the corresponding summation over initial states in the denominator is called the partition function for the initial state. In this case of four atoms each allowed energy is a sum of twelve terms, one for each of the twelve degrees of freedom. This method of calculating equilibrium constants supplants all others for simple molecules and, even for complex molecules, it provides a highly satisfactory method of estimating equilibrium constants.

REACTION RATE THEORY

Closely connected with these ideas is the modern theory of reaction rates. For example, a molecular system in passing from one state of equilibrium to another must pass through a continuous succession of intermediate states. Actually the transition occurs by means of all possible intermediate paths, but the path most economical of energy will be more often traveled. This most frequented path has a configuration of highest potential energy. This configuration is the activated state. The activated configuration is peculiar in that deformation along one degree of freedom, the "reaction coördinate," whether in the forward or backward direction, causes a decrease in the potential energy. In all other degrees of freedom this activated complex is like any ordinary molecule. And so by adding together the energies for each degree of freedom of the activated complex an energy for the complex is obtained. Substituting this energy in the Boltzmann expression one gets the probability of this energy state occurring exactly as for an ordinary stable molecule.

Now when there is equilibrium between initial and final states, half of the activated complexes with a particular energy are rushing along the reaction coördinate from reactants to products, and half are moving with equal speed in the reverse direction. Knowing the energy along the reaction coördinate means knowing the velocity. Accordingly we can calculate the fraction of all complexes which have a particular energy and which are moving forward per unit of length along the reaction coördinate. This fraction is multiplied by their velocity and then by the chance that, having crossed the barrier, they do not return to give the reaction rate constant for this energy. These rate constants for each energy are then added together to get the overall reaction rate constant.

When the above calculation is completed one obtains for v the velocity of reaction,

$$v = \kappa \frac{kT}{h} K^{\ddagger} c_1 c_2 \ldots\ldots \qquad (3)$$

Here the c's stand for the concentrations of the various reactants. $K^{\ddagger}$ is the equilibrium constant between the activated complex and the reactants, that is, it is the ratio of the partition function of the activated state over that for the initial state; $\frac{kT}{h}$ is a frequency which depends only on temperature and for room temperature has the value 5.6×10^{12}; k and h are the Boltzmann constant and Planck's constant respectively; the transmission coefficient κ is the mean value of the chance of a successful completion of the reaction after the system has once crossed the potential barrier.

Equation (3) in its complete generality was first given by the author [2]. Pelzer and Wigner had, however, previously treated certain cases [3]. In particular they had based a numerical treatment of the reaction $H + H_2$ (para) $= H_2$ (ortho) $+ H$ on the potential surface constructed by Eyring and Polanyi [4] who had pointed out that such reactions could be considered as mass points moving on a potential surface in configuration space. Eyring and Polanyi were seeking usable methods of implementing London's approximate calculation of potential barriers. London [5] had recently proposed that quantum mechanics could be used to calculate the potential barriers surmounted in a chemical reaction and had given an approximate formula. Many years earlier Marcelin [6] had proposed a formula which might have become equation (3) had the utility of considering a potential surface in configuration space been appreciated at that time.

Shortly after the present author's paper, Evans and Polanyi [7] sent in their treatment of reaction rates which, in addition to a practically complete formulation of (3), contained the application of (3) to the interpretation of the effect of pressure on rate. Soon after Wynne-Jones and Eyring [8] published their paper on rates couched in the language of thermodynamics and applicable especially to liquids. A part of these and many subsequent applications of (3) have been gathered together in a book devoted to the subject, the theory of reaction rates [9].

Since the formulation of the rate expression given in equation (3) involved only the notion of an activated state in equilibrium with reactants, it is completely general, subject to this equilibrium limitation. Thus we have provided a treatment of all processes involving activation free energies, whether they are chemical or physical. In the latter category such processes are included as viscous or plastic flow of liquids and solids.

We now consider in more detail the nature of the assumptions in the reaction rate theory. The rate of change of reactants to products has been calculated on the assumption that the activated complex is in equilibrium with both products and reactants. We then apply this result when products are absent, assuming their absence does not appreciably influence the forward reaction of reactants to products. A precise, purely theoretical estimate of the approximation made would involve calculating how activated complexes corresponding to reaction in one direction are genetically related to those corresponding to reaction in the other. Such detailed information varies from reaction to reaction, depending on the forces, and is not apt to be available soon. However, there is reason to suppose that in most cases the error made is small. Thus starting without product molecules means that a certain number of hot reactant molecules, that would have been formed at the expense of hot product molecules, are absent in our non-equilibrium system. However, a hot molecule is apt to lose its energy in one of the many ways that does not result in an activated complex and thus be without significant influence on the rate of the process-reactants to products. Guggenheim and Fowler [10] have also discussed this question.

How precisely one should try to estimate the transmission coefficient κ varies with the reaction. Inspection of an approximate potential energy surface usually suffices to tell whether the classical value will be one, a half, or some like value [11]. In such cases the indications are that a quantum mechanical calculation would not significantly change the result [12]. In the unusual case, when a reaction involves a change of multiplicity, a quantum mechanical estimate is necessary. When outside influences cause some atomic or molecular rearrangement to occur more rapidly the following extension of equation (3) is required.

REACTION RATES PROMOTED BY AN EXTERNAL FIELD

If a reactant jumps through a distance λ in the forward direction k' times per second and jumps an equal number of times in the reverse direction, there will be no net motion.

Suppose now a field is applied which does an amount of work W on the system when it moves through a distance of one centimeter in the forward direction. In the distance $\frac{\lambda}{2}$ between a minimum and the crest of the next potential barrier the work $\frac{W\lambda}{2}$ is done. Accordingly systems moving with the field find the barriers lowered by $\frac{W\lambda}{2}$ and raised a like amount for motion against the field. As a result the system will now make the net number of forward jumps per second $n = k'\left(e^{\frac{W\lambda}{2kT}} - e^{-\frac{W\lambda}{2kT}}\right)$ and so have the forward velocity

$$u = n\lambda = \lambda k'\left(e^{\frac{W\lambda}{2kT}} - e^{-\frac{W\lambda}{2kT}}\right) \equiv \lambda k' 2 \sinh \frac{W\lambda}{2kT} \qquad (4).$$

Such external fields can be applied by subjecting matter to shear forces. The use of equation (4) to discuss such cases has led to a theory of viscosity and plasticity [13]. Small electric fields applied to ions allow us to base a theory of conductivity on equation (4). For such fields $2 \sinh \frac{W\lambda}{2kT} = \frac{W\lambda}{kT}$, and the velocity of ions is found to be proportional to the field—Ohm's law. Applied to the large voltage drops across the Helmholtz double layer (4) has been used to build a theory of over-voltage. A slight modification of (4) arises if motions making an angle with the applied potential gradient is explicitly considered.

When $\frac{W\lambda}{kT}$ is very small compared with unity—say about 10^{-7}– as it often is in flow problems—the effect of the field is simply to make the molecules jump $10^7 + 1$ in the forward direction for every 10^7 jumps in the backward direction. Thus by studying what happens in viscous flow one only finds out what molecules are doing in the absence of flow. Consequently any adequate theory of the thermodynamics of liquids must provide the basis for a theory of flow and conversely.

A molecule in the act of jumping from one equilibrium position to the next is essentially gas-like in that degree of freedom. The fluidity will be proportional to the number of such gas-like molecules. As Batschinski pointed out [14], the fluidity, ϕ, of a liquid is proportional to the volume, V, that is, the fluidity is almost independent of a change in temperature, provided it is accompanied by a pressure change that keeps the volume constant. For example, we find $\phi = b\,(V - V_s)$ (5) where V_s can be taken as about the volume of the solid for ordinary liquids and b is a constant. This suggests that the fraction $\frac{V - V_s}{V}$ of the molecules is gas-like and leads to the expectation that the heat capacity associated with motion of the center of gravity of monatomic liquid molecules should be $C_v = \frac{V - V_s}{V} 3 + \frac{V_s}{V} 6$ (6). This formula accounts very well for the specific heat of the liquid rare gases from the melting point to the critical point. The question still remains as to how much extra space must be made available to provide a new equilibrium position for a molecule to jump into. If we assume that the pressure coefficient of viscosity measures the volume required for a new equilibrium position, we find it to be one seventh the molecular volume. If the entropy of melting be assumed to arise from random distribution of the molecules over the new equilibrium positions introduced by the expansion on melting, we find that about one eighth the molecular volume provides a new equilibrium position.

If in accordance with these considerations, a liquid is assumed to possess the fraction $\frac{V - V_s}{V}$ of gas-like degrees of freedom with the rest solid-like, and it is also assumed that extra volume introduces a randomness of position in the manner just described, one can write down a partition function for the simpler molecules which successfully gives the detailed thermodynamics of the three phases and is consistent with their kinetic behavior [15].

PREDICTION OF RATES OF REACTION

If one wants to make diamonds from graphite or sugar from cellulose if is, of course, highly edifying to be told that the theories are all worked out in principle. Most of us, however, would still like to know how actually to carry out the reactions with the idea—needless to say—of contributing any profit to charity. At this stage in the discussion the strictly mathematical chemist or physicist prudently withdraws. The difficulty of proceeding rigorously arises from the same type of mathematical complication encountered in predicting involved stellar events, such as certain eclipses. So far the potential energies of one and two electron systems are the only ones which have been calculated exactly by quantum mechanical methods, whereas the simplest interesting reaction

$$D + H_2 = DH + H \tag{7}$$

involves three electrons.

One is obliged, therefore, to adopt approximate methods or to leave the question alone for the present. Common sense plus even a sidewise glance at the data, however, show a close correlation between the slowness of reaction and the nature of the bonds changed. Thus, there are rate processes which involve predominantly changes in (a) van der Waals', (b) ionic or, (c) homopolar bonds as well as intermediate cases. If the process consists in breaking bonds without simultaneously forming new ones as in dissociation, sublimation, or vaporization, the heat of activation is little if any different from the heat of reaction. Now since new bonds can be formed from such dissociated systems with little or no activation energy, it follows that an upper limit for the activation energy for any process is the heat required to break all the bonds which disappear in the process. However, if other bonds start forming before the initial ones are completely broken the activation energy decreases correspondingly below the maximum value.

Viscous flow and diffusion in condensed phases are typical processes involving the breaking of van der Waals' bonds, and Powell and Eyring [16] found that the free energy of activation for such processes in the liquid state is equal to the heat of vaporization divided by 2.8. These activation free energies are small and such processes are generally very fast.

Water, through its ionizing properties, rapidly catalyzes many otherwise slow double decompositions of inorganic salts. Many organic reactions also involve slow ionization [17]. We shall not consider further reactions of these first two types but discuss rather those reactions which involve the breaking of homopolar bonds.

Most reactions give off heat (are exothermal) in one direction or the other and so absorb heat in the reverse (endothermal) direction. Since the same activated state must be passed through in both directions, the two activation heats differ by just the heat of reaction. Actually inspection of the limited number of bimolecular reactions, such as

$$H_2 + I_2 = 2\,HI \qquad (8)$$

(taken in the exothermal direction) indicates that the activation energy is about one fourth the strength of the bonds broken [18]. For this particular reaction Hirschfelder [19] finds for the heat of activation the relation $\Delta H^* = 0.28\,(D'_{AB} + D'_{CD})$ where D'_{AB} and D'_{CD} are the bond strengths of hydrogen and iodine respectively. An activation energy of 28 per cent of the bonds broken in reaction is also his reasonable, rough estimate for such reactions in general.

Pauling has presented evidence to show that a reaction involving the breaking of homopolar bonds to make new homopolar bonds should be thermoneutral [20]. Stated differently, exothermicity is associated with a change to more polar bonds.

Now since the bonds in the activated complex are intermediate in type between the initial and final bonds, one naturally supposes that in the exothermal direction they will have grown stronger than one would calculate, assuming a thermoneutral reaction. What fraction of the heat of reaction should be subtracted from the 28 per cent of bonds broken to get an improved value for the heat of activation is still uncertain. Probably this fraction, which depends on the polarity of the activated complex, will be very considerably less than half and never large enough to make the activation energy negative.

The value .055 D'_{bc} for the heat of activation seems to account quite well for what is known of atomic reactions of the type

$$D + HCl = DCl + H \tag{9}$$

where D'_{bc} is the strength of the bond broken—in this case HCl.

London [5] was the first to use the perturbation theory of quantum mechanics to obtain a rough formula for the energies of the intermediate states traversed in the course of a reaction. When four atoms are in collision as stated earlier, there are six interatomic distances. With each of these distances is associated a binding energy which has the value and varies with the distance in the same way as it would for that isolated atom pair. Now if such bonds were additive all atoms would tend to form into one large cluster. They do not form such clusters so the bonds are not additive; and spin theory, together with quantum mechanics, shows why. According to approximate perturbation theory each of these six bonds is about 14 per cent coulombic and 86 per cent of the exchange type. The six coulombic bonds are simply additive. The exchange parts of the binding add in a much more complicated fashion. The two exchange bonds present in the two molecules before reaction add up to $\alpha = \alpha_1 + \alpha_2$. The sum of the two final bonds then has the value $\beta = \beta_1 + \beta_2$, and the remaining pair of bonds, having non-zero values only when the four molecules are in collision, may be said to have the sum $\gamma = \gamma_1 + \gamma_2$. London's algebraic formula for the exchange part of the binding can be put into the following simple geometric form [21]: If three vectors equal in length respectively to α, β, and γ are added consecutively, making two 60° angles chosen in the sense to make the fourth side of the resulting quadrilateral short, then this fourth side is the exchange binding for the four-atom system. The rule for addition of exchange binding for three atoms follows immediately if we imagine a four-atom sytem with the fourth atom kept far away. The procedure is exactly as before except that now the three bond energies involving the fourth atom are zero, that is, $\alpha_2 = \beta_2 = \gamma_2 = 0$. A chemical reaction for any three or four atoms thus consists in passing from the initial configurations in which α has its largest value and $\beta = \gamma = 0$ to a state in which only β differs from zero. Clearly in such a process there is no avoiding passing through the configurations for which $\alpha = \beta$. Now if at the same time $\alpha = \beta = \gamma$, the fourth side of the quadrilateral drops to zero so that the exchange binding will drop to zero and only the coulombic binding will be left. From the way exchange binding adds and because of the Boltzmann principle, the critical activated configuration will be one in which γ is as small as possible for the critical values of α and β.

Thus one expects linear three-atom and planar four-atom activated complexes.

Eyring and Polanyi [4] proposed a procedure for calculating these approximate potential surfaces which has become known as the semi-empirical method. One calculates binding energies between atom pairs as a function of interatomic distance, using spectroscopic data in some appropriate equation such as the one proposed by Morse [22]. This binding energy is then assumed to be a certain fraction exchange and the rest coulombic. As already stated, good agreement with experimental rates has been calculated by taking .86 for this fraction for a variety of reactions. This is about the percentage found theoretically [23] for H_2 but theory indicates it should decrease with atoms containing more electrons [24]. Further complications arise when one tries to be precise and consider such things as directed valence [25]. This is a field in which great improvements in both theory and experiment are to be anticipated in the not too distant future. Nevertheless, from these calculated surfaces one can find all the frequencies, moments of inertia, the activation energy, and the transmission coefficients which permit one to calculate reaction rates theoretically Thus one gets a theoretical reaction chemistry qualitatively if not quantitatively coextensive with reality. Historically the development of these potential surfaces, with their emphasis on configuration space, led to the modern statistical-mechanical formulation of rate processes which completes the half-forgotten earlier phase-space formulation of Marcelin and others.

SOME BIOLOGICAL APPLICATIONS

The chemical changes in the living cell are generally catalyzed by enzymes. Without the enzymes much higher temperatures would be required for reaction, bringing with it many unwanted side reactions. One of the processes occurring in a great variety of living cells is oxidation with light emission (luminescence). Recently Harvey has written a book and a subsequent review article which summarize the knowledge of the subject to the time of writing [26].

Some more recent work will be presented here which is of interest for the light it throws on this type of enzyme process and because of its exemplification of general rate theory. There is considerable evidence which indicates that luminescence in organisms as well as in purified material from Cypridina arises from the following reactions [27]:

$$LH_2 + A_n \rightarrow LH_2A_n \qquad (10)$$

$$LH_2A_n + O_2 \rightarrow L + A_n^* + H_2O \qquad (11)$$

$$A_n^* \rightarrow A_n + h\nu \qquad (12)$$

$$A_n \leftrightarrows A_d \qquad (13)$$

Here LH_2 is luciferin, A_n is active luciferase, LH_2A_n is luciferase-luciferin complex. L is oxidized luciferin, A_n^* is excited active luciferase, and A_d is inactive luciferase. Equation (10) is rate determining followed by the rapid oxidation (11) and the rapid light emission by the excited enzyme, luciferase, in step (12). Consequently we can write for the luminescent light intensity, I, the equation:

$$I = sk_{10}\,(LH_2)\,(A_n) \qquad (14)$$

where s is a constant having to do with units and the subscript on k_{10} indicates it applies to equation (10). Now almost all reactions go faster as the temperature is raised, and this is true for the process of equation (10), but there is a complication. The active enzyme, A_n, shifts over to inactive, A_d, and this process of equation (13) is strongly promoted by a rise in temperature. In fact, all luminescent organisms examined increase their luminescence at low temperatures, pass through a maximum intensity, and then become less bright with a further rise in temperature.

If we use A_o for total luciferase, both active and inactive, the following conservation equation is obtained for total luciferase:

$$(A_o) = (A_n) + (A_d) \quad (14). \qquad \text{Also} \frac{(A_d)}{A_n} = K_{13} \quad (15)$$

Substituting (15) in (14) gives $(A_n) = \frac{(A_o)}{1 + K_{13}}$ (16). Finally substituting (16) in (14) gives

$$I = \frac{sk_{10}\,(LH_2)\,(A_o)}{1 + K_{13}} \qquad (17)$$

Equation (17) applies at all temperatures and pressures and leads to interesting results. First we insert in (17) the relationships

$$k_{10} = \kappa \frac{kT}{h} e^{-\frac{\Delta F^{\ddagger}}{RT}} \equiv \kappa \frac{kT}{h} e^{\frac{\Delta S^{\ddagger}}{R}} e^{-\frac{\Delta H^{\ddagger}}{RT}} \qquad (18)$$

$$K_{13} = e^{-\frac{\Delta F}{RT}} \equiv e^{-\frac{\Delta H}{RT}} e^{\frac{\Delta S}{R}} \qquad (19)$$

Now when (17) was fitted to the experimental results of Johnson, Brown, and Marsland for *Photobacterium phosphoreum* by Eyring and Magee [28], the following results were obtained:

$$\Delta H^{\ddagger} = 17{,}220 \text{ cal.}; \quad \Delta H = 55{,}260; \quad \Delta S = 184 \text{ E. U.}$$

The last two values for the heats and entropies for the inactivation reaction of luciferase are very large and so suggest a denaturation process [29].

Equation (17) provides the basis for calculation of how much larger the volume of the activated state of LH_2A_n is than the molecules $LH_2 + A_n$. The value found was $\Delta V^{\ddagger} = 50$ cc. at 0° C. Using (19) the value $\Delta V = 64.6$ cc. at 35° C. was found for the equilibrium between active and inactive catalyst. These results indicate the two contrary effects of pressure on the intensity of luminescence of *Photobacterium phosphoreum.* Thus at temperatures around 10° C. pressures of 500 atmospheres tend to suppress the voluminous activated complex, and so decrease the light intensity to less than half the normal value. While this decrease of the specific rate constant, k_{10}, by pressure persists over the whole temperature range, it is overshadowed in the high temperature range by the stabilizing influence of pressure on active luciferase, preventing its inactivation. On this basis we understand the interesting result that pressure decreases luminescence at low temperatures where $1 >> K_{13}$ and increases it at high temperatures where $1 << K_{13}$. Near the temperature maximum, where $K_{13} = 1$, pressure decreases numerator and denominator in (17) about equally and so has little effect on light intensity.

INHIBITORS OF LUMINESCENCE [30]

Johnson, Eyring, and Williams have examined the effect of various inhibiting substances on light intensity of *Photobacterium phosphoreum.* They systematized all these observations on the assumption that the inhibitors combine with luciferase to render it inactive as a catalyst for luminescence. The inhibitors observed fell into two types.

Type I Inhibitors.

Type I inhibitors combine indifferently with active and inactive luciferase to make two inactive compounds. As a result we have:

$$A_n + rX = A_nX_r \qquad (20)$$

$$A_d + rX = A_dX_r \qquad (21)$$

Here r molecules of inhibitor, X, are combining with active luciferase A_n or inactive A_d to make inactive compounds A_nX_r and A_dX_r respectively. Also the equilibrium constants for equation (20) and (21) are equal, that is, $K_{20} = K_{21}$.

We now write the equations for conservation of total luciferase

$$(A_o) = (A_n) + (A_d) + (A_nX_r) + (A_dX_r) \tag{22}$$

making use of the equilibrium conditions of equations (13), (20) and (21) we have:

$$(A_o) = (A_n) + K_{13}\,(A_n) + (A_n)\,(X)^r K_{20} + (A_n)\,(X)^r K_{20}\,K_{13} \tag{23}$$

or $$(A_n) = \frac{A_o}{1 + K_{13} + K_{20}\,(X)^r + K_{13}\,K_{20}\,(X)^r} = \frac{A_o}{(1 + K_{13})\,(1 + (X)^r K_{20})}$$

From this one expects the inhibited light intensity to be given by the equation

$$I_{24} = \frac{sk_{10}\,(LH_2)\,A_o}{(1 + K_{13})\,(1 + (X)^r K_{20})} \tag{24}$$

The ratio of uninhibited, I, to inhibited, I_{24}, is given by combining equations (17) and (24) to obtain

$$\frac{I}{I_{24}} - 1 = (X)^r K_{20} \tag{25}$$

Now substituting $K_{20} = e^{-\frac{\Delta H_{20}}{RT}}\, e^{\frac{\Delta S_{20}}{R}}$ into (25) and taking logarithms we obtain:

$$\ln\left(\frac{I}{I_{24}} - 1\right) = r \ln X - \frac{\Delta H_{20}}{RT} + \frac{\Delta S_{20}}{R} \tag{26}$$

Plotting $\ln\left(\frac{I}{I_{24}} - 1\right)$ against ln X gives for the slope, r, the value unity for the type I inhibitors examined (that is, sulfanilamide and para-aminobenzoic acid). Thus one molecule of these two type I inhibitors suffices to inactivate an enzyme molecule. For sulfanilamide $\Delta H_{20} = -14{,}400$ calories and $\Delta S_{20} = -38$ E. U., while the para-aminobenzoic acid the values are $\Delta H_{20} = -18{,}000$ calories and $\Delta S_{20} = -50$ E. U.

The negative value of ΔH_{20} corresponds to the fact that increasing the temperature lessens the inhibition. One might, perhaps, have anticipated the experimental result that pressure is almost without effect on the inhibition, as follows:

(1) The inhibitors combine equally well with A_n and A_d and so the addition of inhibitor is not coupled with the large volume increase when A_n goes to A_d.

(2) Most simple reactions involve very small volume changes.

Type II Inhibitors.

Type II inhibitors were explained on the assumption that either A_n or A_d will combine with q molecules of inhibitor Y, to give a substance AY_q. Thus

$$A_d + qY \leftrightarrows AY_q \tag{27}$$

The equation for the conservation of luciferase is then

$$(A_o) = (A_n) + (A_d) + (AY_q) \tag{28}$$

or

$$(A_n) = \frac{(A_o)}{1 + K_{13} + (Y)^q K_{13} K_{27}}$$

Whence

$$I_{29} = \frac{sk_{10}\,(LH_2)\,(A_o)}{1 + K_{13} + (Y)^q K_{13} K_{27}} \tag{29}$$

Combining (29) with (17) we obtain

$$\ln\left\{\left(\frac{I}{I_{29}} - 1\right)\left(1 + \frac{1}{K_{13}}\right)\right\} = q \ln (Y) - \frac{\Delta H_{27}}{RT} + \frac{\Delta S_{27}}{R} \tag{30}$$

A wide variety of compounds are non-specific type II inhibitors and give straight lines when the left side of (30) is plotted against ln (Y) or against $\frac{1}{T}$.

These type II inhibitors reversibly inactivate luciferase by denaturing the protein which is presumably attached to a prosthetic group. The usual protein denaturants, such as the alcohols, acetone, ether, and urethane, apparently belong to type II. For example, approximately three molecules of urethane are found necessary to inactivate an active luciferase molecule. This inactivation occurs with an increase in the heat content of about 55,000 calories and an entropy increase of 182 E. U. However, the inactivation of luciferase in the absence of an inhibitor is about 70,000 calories endothermic and has an increase in entropy of 234 E. U. These energy and entropy values, of course, express the fact that urethane inhibition is greater in the higher temperature range.

McElroy [31] applied the above methods to narcotics, such as phenobarbital, evipal, amytal, barbital, chloretone, chloral hydrate, and trional, and found they are all type II inhibitors. Monoacetin, benzamide, and salicylamide were found to be type I. The barbiturates showed a more complicated behavior. McElroy also pointed out that the type I compounds became relatively more soluble in fat than in water with decreasing temperature, while type II inhibitors showed the opposite behavior. Since at low temperatures substances reveal their bonding proclivities, this indicates that type II inhibitors are strong hydrogen bond formers as was to be expected from their denaturing effect on proteins.

Recently Johnson, Eyring, and Kearns [32] have studied the effects of mixtures of inhibitors on luminescence. Type I with type I or type II with type II inhibitions are readily explained by taking into account all the forms of combination of inhibitors with the enzyme in the conservation equation for luciferase and using the known constants for the separate inhibitors. For a mixture of type I with type II inhibitors it was necessary to assume, in addition, a loose combination between inhibitors. It was possible in this case to choose an appropriate equilibrium constant and systematize these results over an extended temperature composition range.

In summary: the enzyme luciferase can have its prosthetic group poisoned reversibly by a variety of type I inhibitors, notable among which is sulfanilamide. Another group of compounds of type II inhibit by reversibly denaturing the protein attached to the prosthetic group. This is analogous to the reversible denaturation of trypsin. It probably is similar to the reversible liberation of five hydrogen ions by pepsin preliminary to its final irreversible denaturation as observed by Steinhart [33]. LaMer and also Eyring and Stearn have analyzed these latter results [34]. The reversible fivefold ionization of pepsin absorbs 45,100 calories and 127.1 E. U. to give some highly disorganized structure. This process may or may not be followed by the irreversible breaking of some such structure as a cystine bond with an observed energy of 18,400 calories. Luciferase shows an analogous irreversible denaturation which at high temperatures follows the reversible inactivation discussed above.

One expects the results detailed here for luciferase to apply with modifications to a great many of the important enzymes of the living cell. The present treatment, insofar as it requires a conservation equation for luciferase, parallels the early treatment of Langmuir for partly covered surfaces. This does not necessarily indicate, however—as has sometimes been supposed—that surfaces are catalytically important in the cell. The formal treatment of homogeneous and heterogeneous catalysts proceeds along the same lines and much of the analysis outlined above applies to the poisoning of hydrogen catalysts.

The theory of absolute reaction rates as here developed applies to all rate processes in very much the same way that thermodynamics applies to all equilibria. The few examples gathered together give only a sketchy indication of the general applicability of this theory. Certain other examples have been collected in a recent book [9], but most cases remain to be examined.

For References see p. 314.

VIII

STREAMS OF ATOMS

By I. I. RABI

Columbia University

Molecular and atomic beams, in the form of narrow collision-free streams in high vacuum, provide a powerful and simple tool for the study of many important phenomena. These methods were first utilized in a really practical form by O. Stern and his school. The phenomena of space quantization of angular momentum, DeBroglie reflection of atoms and molecules, the Maxwell distribution of atomic and molecular velocities, and the magnetic and electric dipole moments of atoms and molecules were demonstrated and studied by molecular beam methods with very striking and important results. These experiments are now classic in the history of physics.

This discussion will not attempt to present even a small fraction of the applications of this powerful tool but will confine itself to the presentation of some recent developments that have a particular interest in the study of nuclear phenomena.

The atomic nucleus possesses all of the positive charge of the atom and practically all of the mass. It is composed of protons and neutrons confined to a region of about 10^{-12} cm. The protons and neutrons, which can be regarded as the ultimate units of the nuclear structure, each possess a spin of one half of the fundamental unit of angular momentum, that is, $\frac{1}{2}(h/2\pi)$. The fundamental problem of nuclear physics is to gain a quantitative understanding of the forces which are mutual between protons and neutrons, and the nature of their motions within the nucleus. The nuclear problem is very different from what was called the problem of atomic structure, which is concerned with the motions of the circumnuclear electrons. There the physicist starts with a knowledge of the

Coulomb law of forces between electric charges, but for the nucleus this advantage is lacking.

The chief avenue of progress is, presumably, a study of the energies and motions involved in the nuclear system. These experimental data can be obtained partly by a study of nuclear reactions and the masses and energies of the particles involved and secondly in the measurement of other properties which result from the motions, such as angular momentum or spin, and magnetic moments.

Since both the proton and neutron have spin and intrinsic magnetic moment, the total spin of the system will be a combination of these intrinsic spins and the spin introduced by such orbital motions as may be present. The magnetic moment will be a resultant of the combination of the intrinsic moment of the protons and neutrons, as well as that arising from the orbital motions of the protons alone. The measurement of these quantities, therefore, results in very pertinent information about the whole structure. Another important quantity is the electrical quadrupole moment of the nucleus, which is essentially a measure of the departure from spherical symmetry of the electrical charge distribution of the nucleus. In some cases the nucleus may behave as if it were an oblate or prolate charge distribution about the axis of spin. The connection between such distribution and the nuclear forces and motions is close.

MOLECULAR BEAM METHODS

The application of molecular beams to these problems arises from the simple experimental situation that can be obtained. A beam of atoms or molecules is readily obtained. In a space, exhausted of gas by means of a vacuum pump, is placed a vessel which is termed an oven. In the oven some material is placed and heated till a sufficient vapor pressure develops. A gas or vapor is equally good. The molecules then stream out in all directions from a slit or orifice due to their natural heat motion. A second slit intercepts all but a narrow pencil of these molecules or atoms. Beyond this point there is a molecular beam, the components of which are atoms or molecules moving in straight lines with a velocity distribution appropriate to the temperature of the oven, and free of collision if the vacuum is high enough. To complete the apparatus there is a detector for the beam, which may be movable in order to explore its form. This is clearly a very elementary situation. For example, if the molecules have a magnetic moment, by placing a sufficiently strong magnet to one side of the beam, they can be deflected toward or away from that side, depending on the orientation of the atomic or molecular magnet. The deflection may be measured by moving the detector an appropriate distance. From the magnitude of the deflection the moment can be evaluated. Indeed the first measurement of the magnetic moment of the proton by Stern, Estermann, and Frisch was done in essentially this manner. The results obtained are very direct and striking.

Although the principle of the deflection method of measuring magnetic moment is very simple, the evaluation of the experimental results is rather difficult. This is due to the fact that the atomic or molecular beam is not one in which all the particles have the same velocity; they are distributed more or less according to Maxwell's law of velocity distribution. One can take account of this distribution, but the beam departs very considerably from the law as a result of collisions with the molecules of the residual gas in the vacuum chamber. Another very serious difficulty is associated with the fact that the force exerted on the atomic magnet is proportional to the departure of the magnetic field from uniformity. The accurate measurement of this departure has always been a very difficult operation.

Accordingly methods have been developed in recent years that measure the magnetic moment and do not involve either the velocity distribution or the inhomogeneous field which tends to deflect the beam. These methods now go under the general heading of "radiofrequency spectroscopy." They are more allied to the older spectroscopic measurements of the Zeeman effect than to the deflection methods. The molecular or atomic beam provides the simple physical environment.

The fundamental principle is easily seen from an elementary mechanical analogy. An atom possessing spin and a magnetic moment is placed in a uniform magnetic field. Because of its spin it will behave like a gyroscope, and the torque which the magnetic field exerts on the atomic magnet to line it up in the direction of the field will not result in this alignment. Instead, because of the gyroscopic effect, the atom will precess about the direction of the magnetic field with its spin axis at a fixed angle with respect to the magnetic field. The frequency of this precession is given by the simple relation from elementary mechanics,

$$\nu = \mu \frac{H}{Jh} \qquad (1)$$

In this formula ν is the frequency of the precession, μ is the value of the magnetic moment of the atomic magnet, H the magnetic field, and J is the magnitude of the spin expressed in units of $h/2\pi$—the natural unit of spin according to the quantum theory. Finally, h is the familiar Planck's constant that pervades all of twentieth-century physics. The magnetic field H involved in this formula is a uniform field. The spin J is generally known from collateral evidence or can be measured by other methods. According to the quantum theory it can only have the values 0, ½, 1, 3/2, etc.

Considering the choice of J as known or given by some other experiment, it is clear that a measurement of ν, the precession frequency, is all that is necessary. Since the precession is not observable directly, as in a top or laboratory gyroscope, recourse must be had to an indirect method. An oscillating magnetic field, much smaller in magnitude than the constant uniform field, is superimposed in a perpendicular direction. The effect of this field is to apply a small but rapidly variable torque on the atomic magnet. If the oscillating field is very small its effect will generally be nil over a period of time. However if the frequency of oscillation is close to the frequency of precession, ν, then a resonance will occur, and the effect of the field will pile up with each oscillation instead of averaging out to zero. The physical result will be a change in the angle of the spin axis with respect to the uniform field.

This effect is easily detectable, because the angle determines the projection of the nuclear magnet in the direction of the field and this, in turn, determines the deflecting force which a non-uniform field exerts on the atom. When the force changes, the atom or molecule is deflected to a different place, and causes a change in the number of atoms which arrive per second at the detector. The number of atoms which arrive at the detector is measured, therefore, as a function of the frequency of the oscillating field. This number is a constant except in the region of frequency in the neighborhood of ν as given by the equation. A plot with frequency as abscissa reveals a straight horizontal line with dips in these regions. Thus these spectral frequencies can easily be measured to one part in twenty thousand.

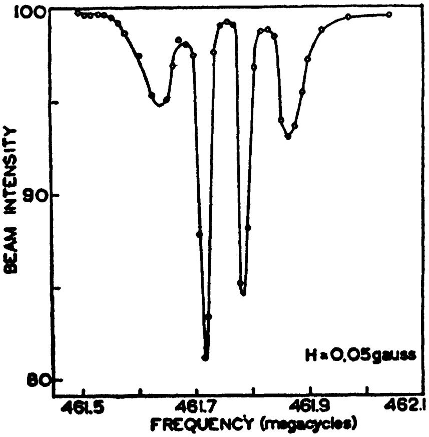
If molecules are the chief constituent of the beam each atomic nucleus reveals its own spectrum independently, except for slight interactions between the nuclei. If the beam consists of a mixture of different atoms, such as a mixture of sodium and potassium or the different isotopes of the same element, the above statement still holds true. The characteristic frequencies of each atom appear in their appropriate place.

CONNECTION WITH SPECTROSCOPY

Another view of the process described above, and one that is more general than the simple mechanical analogy, is based on the principles of the quantum theory.

An atom or molecule in a magnetic field is capable of existing in one of a number of different energy levels. It is capable of passing from a level of high energy to one of lower energy by the emission of radiation according to the well-known Bohr relation. This relation states that the frequency of the radiation multiplied by Planck's constant h is equal to the difference in energy between the two states. This process can also be made to proceed by illuminating the atom or molecule with the same radiation. Conversely, the atom or molecule can be made to go from a lower to a higher energy level by this illumination, because the atom can absorb the energy of the radiation, provided the frequency is appropriate for the Bohr relation. The oscillating magnetic field according to Maxwell's theory, is a radiation field and provides the low frequency radiations that are necessary to cause these transitions in exactly the same way as the absorption of the sun's radiation causes the Frauenhofer lines in the solar spectrum. The same laws hold in the case of visible radiations where the frequencies of the order of 10^{15} cycles per second as in radiofrequency spectroscopy, with frequencies of the order of 10^5 cycles per second.

The low frequencies concerned in radiofrequency spectroscopy imply that the energy levels are very close together. This method, therefore, is a tool for measuring very delicate and minute details of atomic and molecular structure. Indeed, a whole complicated radiofrequency spectrum can lie in a frequency interval less than the natural width of a very fine optical spectral line.



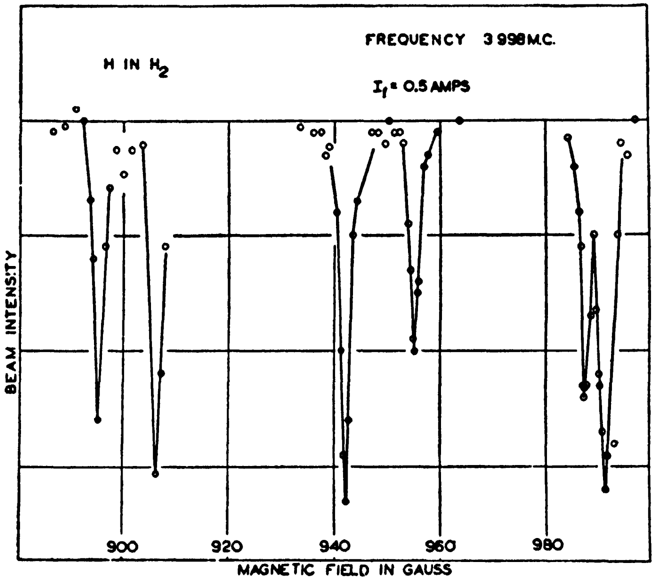
Curve A. In this figure is shown the radiofrequency spectrum of potassium, atomic weight 39. The four lines represent the Zeeman effect of this spectrum in a field of 0.05 gauss. The spectrum lies in the region of frequency of 461.7 megacycles. The fine structure of the main lines arising from the Zeeman effect cannot be observed at such low fields by optical methods, since the resolutions there obtainable are insufficient by a factor of more than 1000.

The curve is obtained by observing the intensity of the beam as a function of the frequency of an oscillating magnetic field.

RESULTS

In the preceding paragraphs, methods of obtaining radiofrequency spectra have been described. In certain simple cases the interpretation is very direct, in others a large number of lines are involved. If one is dealing with an atom which has a nuclear moment but no resultant electronic moment, as in the case of barium, there is a single line for each isotope. The frequency at which this line occurs gives the precession frequency of the nucleus in the magnetic field, and the evaluation of the nuclear moment follows immediately from equation 1 above.

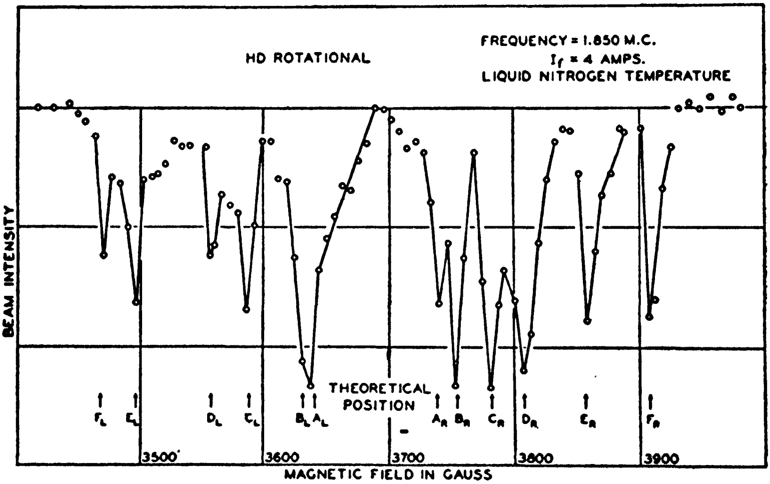
A more complicated situation arises in dealing with a molecule, even as simple as that of hydrogen, H_2. Here the effect of electronic spin may be neglected, because the two electrons in H_2 are oppositely oriented and cancel out one another.



CURVE B. This figure represents the radiofrequency of the H_2 molecule. It is obtained by plotting the intensity at a fixed frequency of the oscillating magnetic field and varying the magnitude of the uniform magnetic field. This procedure is equivalent to varying the frequency and is sometimes more convenient. The locations of the lines give all the necessary data for the physical conclusions described in the text.

However, in addition to the external magnetic field, which causes the two protons to precess, there are the magnetic fields produced by the protons themselves which, because of their spins, are equivalent to small magnets. An added complexity is the magnetic field produced by the rotation of the molecule as a whole. Since the constituents of the molecule are electrical, this rotation produces a magnetic field at the positions of the protons which must be added to the external field and the field produced by the electronic magnets. The magnitude of these fields depends on the internuclear distance, the rotational state of the molecule, and the details of the molecular structure.

From a complete analysis of the radiofrequency spectrum combined with the theory of the H_2 molecule, it is possible to evaluate the proton moment from (1) its rate of precession in the external magnetic field and (2) the field produced by one proton at the other, since we know the internuclear distance from band spectra measurements. In addition the field produced by the rotation is measured. The value of the proton moment may be checked by measurements on the hydrogen-deuterium molecule, HD, in the state of zero rotation. In this state there are no disturbing effects from the interaction of the two nuclei with one another and, of course, no rotational magnetism. The spectrum consists of two lines: one of the precession frequency of the proton, the other that of the deuteron. The results are in accord within the limits of experimental error.



Curve C. This figure represents a portion of the radiofrequency spectrum of HD and was taken in the same way as the H_2 spectrum shown in Curve B. The lines below, labeled "theoretical position," show how closely the spectrum can be accounted for with the aid of the theory.

Similar experiments were performed on the deuterium molecule, D_2. These experiments led to the discovery of the quadrupole electric moment of the deuteron, that is, the deuteron is not spherical in shape but may be more correctly regarded as an ellipsoidal figure spinning about its longer axis. The following considerations and experimental facts lead to this inference. Since the structure of the D_2 molecule is known, as well as the internuclear distance, we can predict from a knowledge of the deuteron magnetic moment obtained from the HD experiment the spectrum which should be observed. This prediction makes the tacit assumption that no quadrupole electric moment is present. However the experiment showed that the separation of the lines in the D_2 spectrum was greater by a factor of six than the predicted position. This departure is caused by the fact that, in addition to the magnetic interactions described above, there is an electrical interaction between the deuteron and the other electrical charges in the D_2 molecule. Because of the departure of the form of the deuteron from spherical symmetry the magnitude of the energy of this interaction depends on the orientation of the deuteron axis with respect to the rest of the D_2 molecule. The magnitude of this interaction is obtained from the analysis of the radiofrequency spectrum. From the theory of the deuterium molecule the electric field exerted on the deuteron is calculated, and the magnitude of the quadrupole moment evaluated.

The existence of a quadrupole electric moment in the deuteron is of profound significance. It can be shown that, if the force acting between the proton and neutron is "central" in nature, that is, if the force depends only on the distance between the particles and not on the relative orientation of their spins and the line joining proton to neutron, then the deuteron cannot possess a quadrupole moment. The charge distribution must be spherically symmetrical. The existence of the quadrupole moment shows that the force is not of the simple type of gravitational and Coulomb forces, but depends also on the spin orientation. The suggestion has been made that these forces can be correlated with the emission of mesotrons from nuclei in a manner similar to the correlation between the Coulomb force between charged particles and the emission of light quanta. However no completely consistent account of the nature of nuclear forces has yet been given.

We have indicated in these few paragraphs the kind of problem that can be dealt with by the methods of radiofrequency spectroscopy. There still remain large fields in the domain of nuclear and molecular structure as yet untouched. What has been done with oscillating magnetic fields can also be done with oscillating electric fields, and thereby an entirely different domain of phenomena opened to investigation. Another fertile field lies in the investigation of artificially radioactive substances, where the quantities available are too small for ordinary spectroscopic study. The experiments described above represent but the first step in a new and enticing field of experimental physics.

For References see p. 316.

IX

ADVENTURES IN VACUUM CHEMISTRY

By K. C. D. HICKMAN

Distillation Products, Inc., Rochester, New York

THESE Sigma Xi National Lectureships are an adventure in coöperative science where many topics are brought before groups composed of many branches of learning. The adventure is important because discoveries arise where knowledge is blended; events occur at interfaces. High-vacuum chemistry is a typical adventure in blended science. I shall consider two contributing themes: the high-vacuum still itself and the researches, chiefly concerned with oils and vitamins, which are being done with the still. In describing these researches I shall ask your indulgence while I recount chiefly the work of my colleagues at our laboratory for it is about their work that I am most recently and fully informed.

All human experience is complicated by the presence of air. Because the atmosphere gives life to organic nature, its presence has been considered normal, even in the scientific laboratory. Investigations made in the laboratory, however, are all modified and many are entirely inhibited by the presence of air. The scientist has been like an explorer in the desert who wishes to taste a new kind of game. He cooks it, cuts a slice, and puts it into his mouth; but what he tastes is not meat but the layers of sand that have settled on either side. If he wants to know what the meat really tastes like, he must go into a tent, exclude the desert storm, and cut a fresh piece. Vacuum chambers are areas from which molecular areas can be excluded while the phenomena are studied and our data revised. Chief amongst these phenomena is distillation.

Why high-vacuum distillation? For the benefit of those who are forgetful of their kinetic theory of gases and the physics of distillation, I shall deal with the matter pictorially.

Communication No. 62 from the Laboratories of Distillation Products, Inc.

All molecules are in vigorous thermal motion, bumping against one another, the larger molecules traveling more slowly than the smaller, according to the square roots of their molecular weights. Now, in a liquid—and I refer you to Section A of Figure 60, where you see a liquid resting in the half-completed diagram of a flask—a molecule in the position *a* in the center of the flask is attracted electrically and gravitationally on all sides. It is, therefore, within the limits of collision, free to travel in any direction. The molecule *b* in the surface of the liquid is in a very different situation. It is being attracted into the liquid from behind but there is no attractive force in front. It therefore finds it very difficult to leave the surface and extra energy must be put into it to make it do so. This extra energy can be acquired by merely waiting for an unusually hard collision, or heat can be put into the liquid to accelerate the over-all thermal velocity of the group and thus make it easier for one of the molecules in the surface to leave. But whatever happens, the molecule that leaves is hotter than the average; that is to say, by definition, distillation is done by heat; again, by the application of vigorous collisions. But large, long chain organic molecules are likely to be damaged by too vigorous agitation and are liable to break in the middle. *High-temperature distillation involves thermal decomposition.*

Suppose that spontaneously, or by the application of external heat, molecule *b* does escape, and in a perfect vacuum travels to a sticky place *c*, to which it adheres. By "sticky" we mean, of course, cold so that it cannot reëvaporate. Under these circumstances, distillation has been done in the simplest possible manner by a single molecular transfer. The molecule has left the parent liquid once and traveled without collision to the condenser. This form of distillation has been defined as *molecular* distillation [1] and it occurs at the lowest possible temperature. Suppose, however, as in Figure 60, Section B, the surface of the liquid is covered with a layer of foreign gas (air) indicated by the collection of crosses. Now the molecule *b* emerges from the liquid, hits an air molecule, and is knocked back into the liquid. It may stay near the surface a moment and be ejected again, only to have the same difficulty and eventually be brought back into the liquid. When it has entered the liquid it may wander completely away from the surface and another molecule must make its try to get out. So air prevents evaporation and causes a molecule to make many attempts to escape. Suppose, however, that we complete the geometry of our flask and attach it to a condenser, as in Figure 60, Section C. Again let all the air be removed, and again let molecule *b* emerge from the liquid. You will see that there is only about one path that it could take to arrive in the condenser without the danger of being sent back into the liquid, and that is the path shown by the dotted lines. Actually the molecule is likely to follow any one of the other paths and eventually arrive back in the liquid. The chance of its emerging and condensing at exactly one strike is remote indeed. Therefore to do distillation at the lowest possible temperature two things must be done—*remove the atmosphere* and *remove the flask itself.*

THE SCOPE OF SHORT-PATH DISTILLATION

The modern short-path still is the chemical engineer's solution to these drastic requirements. Before considering the construction and behavior of the still, let us examine what it can do [2] [3]. Quite simply it renders distillable a whole category of substances, natural and synthetic, which could not be distilled previously by any known means. The category is large—all the "fixed" fats and oils, once called fixed because they would not volatilize unchanged; the natural waxes, resins, vitamins, sterols, and many synthetic polymers and drugs of high molecular weight. To pick an example, there are several million pounds of high molecular weight plasticisers and polymer intermediates produced annually in America alone which could be purified by molecular distillation should the need arise. In the molecular still they will evaporate unharmed at about 150° C. lower than they would boil (with decomposition) in the ordinary vacuum still with separate boiler and condenser. The range of usefulness of the short-path still is for organic chemicals of molecular weights between 250 and 1,200. If much oxygen, sulfur, or halogens are present in the molecules, the range is considerably lower.

The implied revolution in fat technology has not transpired to any extent except among the oils which are rich in vitamin A, such as shark liver and pollack oils, and it is these which have carried the development of short-path distillation. The first molecular still for organic substances was described in England by C. R. Burch [4] who is indeed the father of the industry. The industry itself, however, with its collateral research has grown up in Rochester, New York, where concentrates of vitamin A are produced [5] [6]. The focus of our story will be this vitamin industry.

EVOLUTION OF THE MOLECULAR STILL

The simplest practical interpretation of the molecular still is the flat evaporator shown in Figure 60, Section D. Here a layer of liquid is held in the bottom of a flat glass flask. Heat is applied underneath, a vacuum is applied to the space between the floor and the ceiling, and soon molecules begin to evaporate; most of them—60 per cent or 70 per cent, that is to say—collect on the ceiling and drip into the receiver. Thus we would appear to have satisfied all the requirements for a good molecular still.

We have, however, failed to consider one very important factor, namely, the surface of the liquid that is distilling. Remember that it is only this surface that is involved in distillation. Momentarily, none of the other molecules in the still need be there. Their only reason for being in the still is so that they may come to the surface when wanted and take their chance at distillation. Where we are dealing with a mixture of many substances all of different volatilities, we have a mixture of molecules with statistically different chances at any moment that they will receive enough energy to distill. Our purpose, then, in effecting distillation is to bring a molecule to the surface and ask it the all-important question: "Do you wish to distill?" And it is in the statistically varying answers given by the molecules that we can effect separations. But whereas only the molecules that are in the surface are needed for distillation, all the molecules in the still are held there at the same temperature as the surface and are, therefore, exposed to conditions of decomposition. You can see that if the still were ten times deeper we would have ten times as many molecules exposed to decomposition. If, on the other hand, we could shorten the still a hundred times, we would prevent decomposition one hundredfold.

There is another important reason for having a very thin layer of distilland, and that is the lack of ebullition in the molecular still. Consider the teakettle on the kitchen stove. After water has been put on to boil a singing noise is produced which, in turn, changes to a sharp hiss as the kettle boils. These sounds are caused by the generation of bubbles of steam and dissolved air which eventually come to the surface and burst, causing violent agitation of the contents (Fig. 60, Sec. E). We may say with confidence that every molecule in the kettle has the opportunity of reaching the surface many times a minute and taking its chance of distillation. If we had placed a mixture of alcohol and water in the kettle instead of plain water, we could be certain that every one of those alcohol molecules would have ample opportunity of being brought to the surface and distilling preferentially from the less volatile water. In the pot still of Figure 60, Section D, however, the conditions are vastly different. There must be no bubbles of gas underneath the surface of the liquid, for if this liquid were a centimeter deep then the formation of a bubble would entail the generation of vapor at approximately 1 mm. mercury pressure. But we are trying to generate vapor not at a millimeter but at 0.001 of a millimeter. Therefore we must heat the contents so gently that there is no ebullition. Coupled with the absence of ebullition is the fact that we are dealing with long, large molecules of high molecular weight. The solution itself is very viscous, and it is a matter then of practical experience that the interchange between the surface and the body of the liquid is so poor that many molecules that should come to the surface to be asked that question, "Do you wish to distill?" do not actually arrive during the course of hours. A distillation designed to protect materials from thermal decomposition may actually entail their decomposition, because

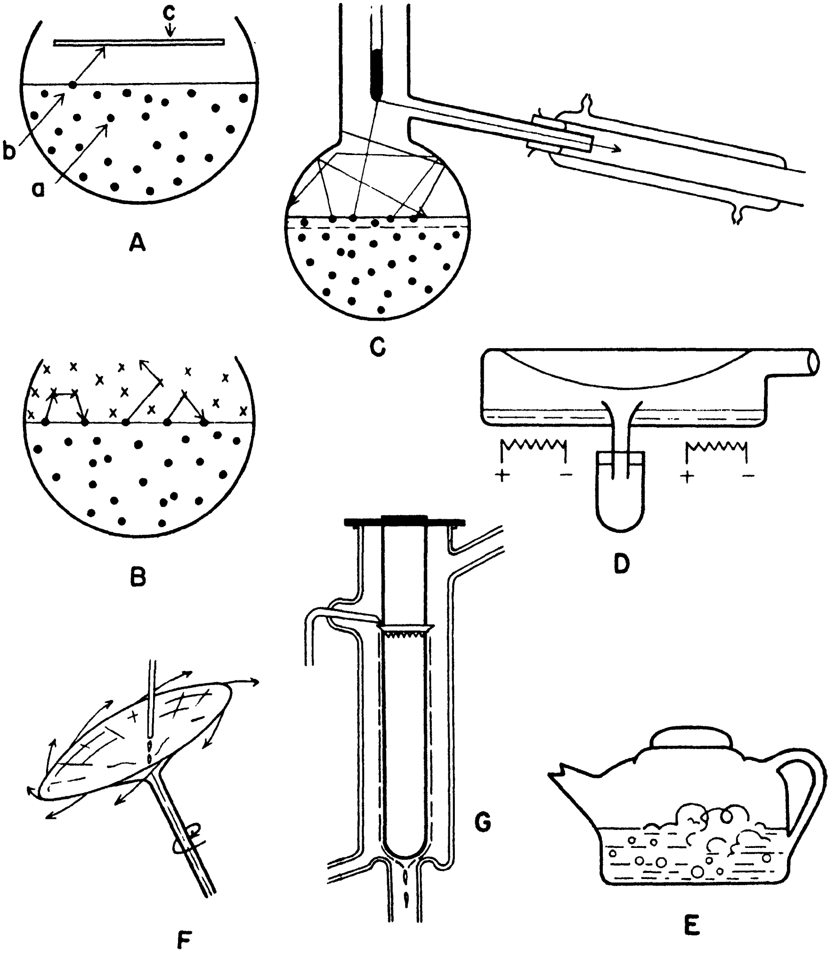


Fig. 60. Elementary physics of distillation.

the mixture must be kept at the high temperature of distillation for too long a time. When molecular distillation was first formalized [7], the physicists advanced calculations to show that exposing viscous organic liquids in trays at the depth of a centimeter or so would give ample opportunity for renewal of the distilling surface; that the composition at the surface at any moment would be representative of the composition of the bulk of the liquid. Sad experience, however, has shown that this is not generally the case. The commonest experience to the contrary is the resistance which liquids offer to *degassing* under vacuum. When an organic mixture is placed in the molecular still, three reserves of gas have to be drawn away by the pumps in producing a vacuum. First and easiest is the residual air in the vessel. Next comes the air dissolved in the liquid, and finally the semi-volatiles which are evolved when the liquid is heated but are too volatile to recondense on the condenser. They may cruise around in the vacuum space and greatly hinder distillation.

Let us consider first the dissolved air. The air may be regarded as a distilling constituent which boils at, say, 500° C. below the rest of liquid. With such a difference in boiling point substantially every molecule of air brought to the surface will avail itself of the opportunity to evaporate, so that if a monomolecular layer of oil containing dissolved air were exposed in a molecular still, then in the course of a single set of molecular collisions, lasting perhaps a microsecond, all the oxygen and nitrogen molecules would leave. Now let us study how long it actually takes to remove all dissolved air from thicker layers of oil. Suppose cod-liver oil is placed in flask C of Figure 60 and vacuum is applied; that the flask be gently warmed to, say, 80° C. Now let readings be taken by a sensitive manometer such as the Pirani gauge. Suppose also that the vacuum pumps possess the inherent capacity to produce a vacuum of 10^{-4} mm. We shall find that during a day the pressure will gradually fall to about 0.1 mm.; leaving the flask overnight, the pressure may have fallen to 0.01 mm. by morning. If at this stage we pick up the apparatus and shake it, a vigorous evolution of bubbles will be observed, and the pressure will probably rise to about 1 mm. In other words, although the surface of the liquid is emitting gas so slowly that the vaccum space under steady pumping contains 0.01 mm., the saturation pressure within the liquid is perhaps a hundred times as great. If the same experiment is done with the apparatus E of Figure 60, it will be found that a pressure 10^{-3} mm. will be establishd in perhaps three hours, and the pressure will not rise to more than 10^{-2} mm. when the apparatus is shaken. If it is not shaken a molecular distillation can be allowed to proceed. Suppose, however, the oil were placed in a falling-film still of the kind shown in Figure 60 G. Three passages over the evaporator, involving a total of about a minute exposure for each element of the distilland, at 80° C. will remove enough of the dissolved gas for distillation to proceed. Again, let the oil be passed over a high-speed centrifugal evaporator (Fig. 60 F). We shall find that in one pass all the dissolved gas is removed to equilibrium so that if the liquid be allowed to flow into a closed flask attached to a manometer, shaking the liquid will not release any more gas. You will see, therefore, that there is every reason to reduce the thickness of the liquid,

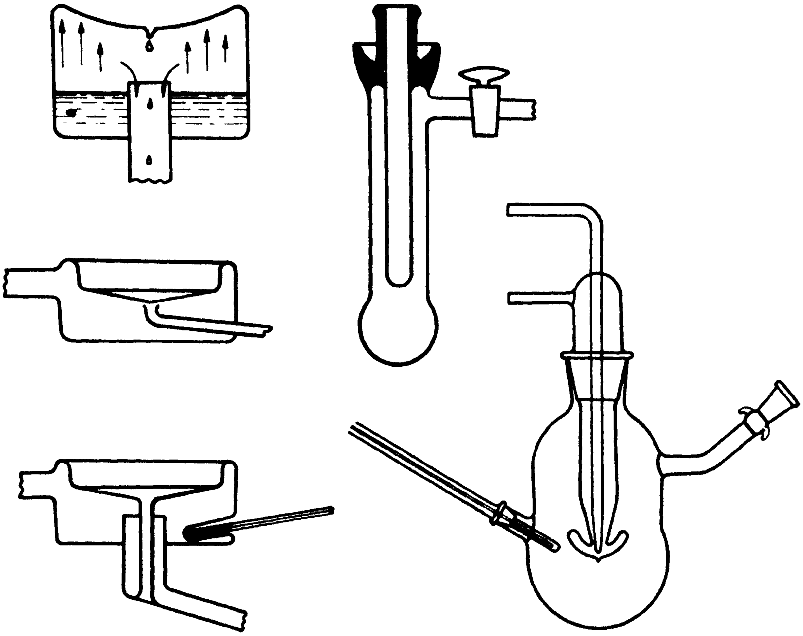


Fig. 61. Miscellaneous types of glass pot stills.

because this evolution of dissolved gas is an excellent index of the availability of the surface to dissolved molecules.

One of the chief problems with which we have been preoccupied during the past ten years has been reducing the distilland layer to a thin and amenable condition for distillation of mixtures, and I shall describe some of this work now. Figure 61, which is largely self-explanatory, shows various pot stills which have been devised in different labora-

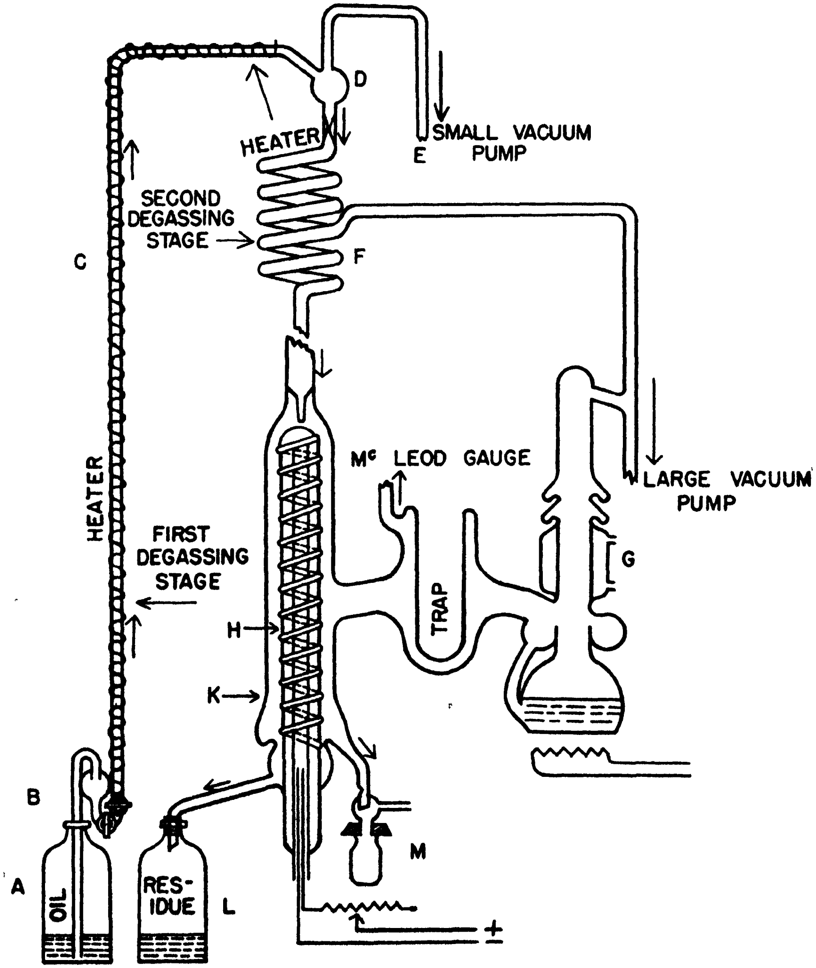


Fig. 62. Glass falling-film still.

tories for evaporating material of high-boiling point. Figure 62 shows the falling-film still, which is an important step in the development of high-vacuum distillation.

THE FALLING-FILM STILL

The oil to be distilled, known as the distilland, is held in a vessel A from which it passes by way of a drop counter or other measuring device B, through a metering stopcock up a warmed tube C, and into an "explosion bulb" D where much of the dissolved gas is given off to a separate small vacuum pump E. The distilland then passes down through a constriction into an electrically heated spiral F, on the bottom surface of which it flows and gives up more gas to the fore-vacuum lead of the condensation pump G. The distilland then passes through a constriction onto the falling-film column H. This is an upright closed cylinder or pole which is heated internally by electricity or by fluid. In the example shown it is warmed by a vapor such as ethyl lactate. A pattern is embossed on the pillar so that the oil falling out of the top spreads out into a more or less even film and ripples its way down to the bottom where it passes to the residue bottle. Surrounding the column is a glass jacket K conveniently cooled by the draught from a small fan. Vacuum is produced by means of the vapor pump G and, if necessary, a refrigerant trap filled with ice or dry ice. The distillate is collected in the receiver M. After one pass has been made the apparatus is shut down, the residue bottle L and the incoming oil bottle are interchanged, and the oil is sent through again to be distilled at a higher temperature. If there are X constituents in the oil to be separated it is necessary to perform at least X, and in practice many times X, number of separate passes through the still, each accompanied by a separate change of receiver.

The falling-film still spreads the distilland in a thinner and more turbulent layer than any previous evaporator, but even so the apparatus is by no means perfect. Its chief defect is the tendency of the oil to gather into local streams as it flows over the column. The trace of oil left on the bare spots becomes too hot, the rest of the oil receives less than the required opportunity to distill. This behavior is an inherent property of the surface still. Liquids, in general, exhibit a diminishing surface tension as they become hotter. Now the molecular still demands that a layer of hot oil shall be facing a cooled condenser and it is a matter of practical experience that much more heat is lost by radiation [8] than is spent in evaporation. There is thus a steep thermal gradient through the oil between the hot evaporator and the cold condenser. The surface tension of the oil in contact with the evaporator is lower than the tension of the outside, so the oil tends to "wrap itself up" away from instead of toward the column. This means that the liquid film is in unstable equilibrium and will tend to fall in rivulets. Such rivulets are familiar to anybody who has spilled fat or water on a hot stove and watched the liquid gather itself into globules. Now there are two ways of dealing with unstable equilibrium. One is to balance the system so perfectly that it is unlikely to tumble either way and the other is to alter the conditions so that there will be no tendency to tumble, that is, bring the system into stable equilibrium. Arguing along the first line, it was thought that if oil were distributed *perfectly* at the top of the column this separation into rivulets would not occur. To test this a rotary distributor, like a lawn sprinkler, was added to the top of a falling-film column. There was no improvement. However, having put a rotating shaft into high vacuum without serious leakage, it appeared possible to rotate the still rather than the distributor.

ADVENTURES IN CENTRIFUGAL DISTILLATION

A rapidly rotating, centrifugal, high-vacuum still would hold the distilland under the constraint of centrifugal force which, if necessary, could be made many hundred times that of gravity. The oil would be pressed against the surface so vigorously that it would have to stay in place. A small experimental still, finished in 1935, is shown in Figure 63. Here a slightly conical rotor with a deeper cone at the center, spun out of sheet aluminum, was mounted on a shaft which in turn was placed in bearings housed in a vertical backplate. The various leads needed to service the still were brought through the plate and the whole assembly was covered with a bell jar (removed in photograph). A removable condenser could be placed in front of the rotating cone. Oil was allowed to fall through a glass tube to the center of the cone which was rotated at 3,000 to 5,000 revolutions a minute.

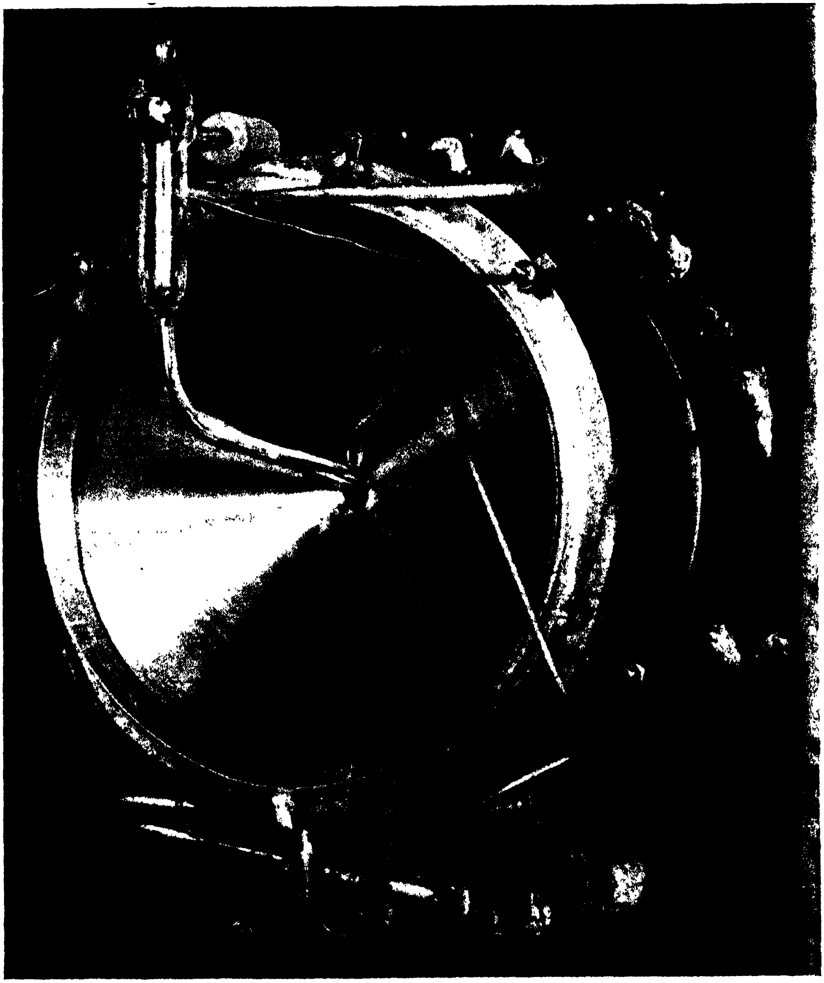


FIG. 63. Early centrifugal still—bell jar removed.

The oil passed across the surface and was thrown off as a fine spray from the edge into a water-cooled collecting gutter. By putting on intermittent streams of colored and colorless oil it was possible to find out by watching the rings of color pass outward how quickly the oil went across the surface. We were interested to note that the total time of passage in this 7-inch diameter still was less than a tenth of a second. When the oil was kept very hot and under distillation conditions, the actual travel across the distilling zone was of the order of one-fiftieth of a second. Now if you know the area of a surface, the amount of material you are putting on the surface, and the time of flow across the surface, the thickness can be determined by a simple proportionality sum. It was at once apparent that the thickness of the oil was varying from a tenth of a millimeter to less than a hundredth of a millimeter. Under these conditions it was evident that the time of exposure had been reduced from the few hours of the glass pot still nearly half a million times in our first excursion among centrifugal stills. As a matter of interest a comparison of times of exposure and film thicknesses is shown in Table I, and in Table II a comparison of probable thermal exposures. It is assumed that every rise of 10° C. will double the thermal hazard to the average distilland. It is evident that the small rotating still provides a reserve of thermal stability which may be more than the

TABLE I

Transition from Pot Still to Centrifugal Still

Approximate Date	*Still*	*Approximate* Distilland Thickness*	*Approximate† Molecular Thickness*	*Approximate Time of Exposure*
1922	Laboratory pot still (Fig. 60 D)	1–5 cm.	5×10^7	1–5 hr.
1935	Industrial falling-film still	1–3 mm.	5×10^5	2–10 min.
1930	Laboratory falling-film still	0.1–0.3 mm.	5×10^4	10–50 sec.
1940	Industrial centrifugal still	0.03–0.06 mm.	10,000	0.1–1 sec.
1936	Laboratory centrifugal still	0.01–0.02 mm.	3,000	0.04–0.08 sec.
1942	High-speed centrifugal rim still	0.001–0.005 mm.	400	0.001–0.005 sec.

* Assuming similar throughput for same unit area of all stills.

† Assuming that the molecule of glyceride fat has an effective diameter of 15 Å.

margin needed by the distilland. With so much time reserve in hand, thermal liberties can be taken with the distilland, and the temperature raised far above the usual point of decomposition. Thus melted butter can be put into the still and turned into a gas of butter moving outward with average molecular speed of about 400 feet per second at 300° C., and a saturation pressure of 5 or 10 microns. This butter stream contains so much energy that a light pendulum hung in front of the evaporator will lie horizontal in the stream. A small fan, such as a drugstore fan with the motor removed, will rotate so rapidly that it exceeds its bursting strength and flies to pieces, shattering the bell jar. We do not recommend this experiment, but it is indicative of the magnitudes involved.

TABLE II

Relative Thermal Exposures

Pressure	*Temperature**	*Relative Thermal Coeff.†*	*Time*	*Relative Decomposition Hazard*	*Kind of Still*
Atmospheric	360° C.	2^{23}	1 hr.	1.5 trillion	Simple flask
10 mm.Hg.	270	2^{14}	1 hr.	3 billion	Claisen flask
1 mm.	220	2^{9}	1 hr.	92 million	Wide-necked pot still
1 mm.	220	2^{9}	1 min.	1.5 million	Petroleum flash still
001 mm.	130	1	1 hr.	180,000	Molecular pot still
001 mm.	130	1	1 min.	3,000	Molecular falling-film still
001 mm.	130	1	1 sec.	50	Molecular centrifugal still
.001 mm.	130	1	.02 sec.	1	Molecular centrifugal still

* Boiling points of dibenzylphthalate, for example [9].
† Assuming that hazard doubles with each 10° C. rise.

Another kind of rotating still is shown in the photograph (Fig. 64) which was taken with the bell jar and condenser removed. In this model there is a minor change in design, for the spent distilland is "picked up" at the edge of the rotor instead of passing to a water-cooled gutter. We do not yet know how to measure the temperature of the distilland with any great accuracy as it flows across the evaporating surface. If a thermocouple is placed in contact with the rotating plate most of the oil is splashed off in a violent spray, and the friction raises the thermocouple above its proper temperature. Instead, it is better to have a bifilar pendulum hanging in front of the evaporator. The oil vapor tends to thrust the pendulum outward. If the oil is evaporating too fast near the center the pendulum will twist. Similarly, if it is evaporating too fast near the edge it will twist in the opposite direction. But if the heaters are adjusted properly, the entire pendulum swings out a distance which is indicative of the rate of evaporation.

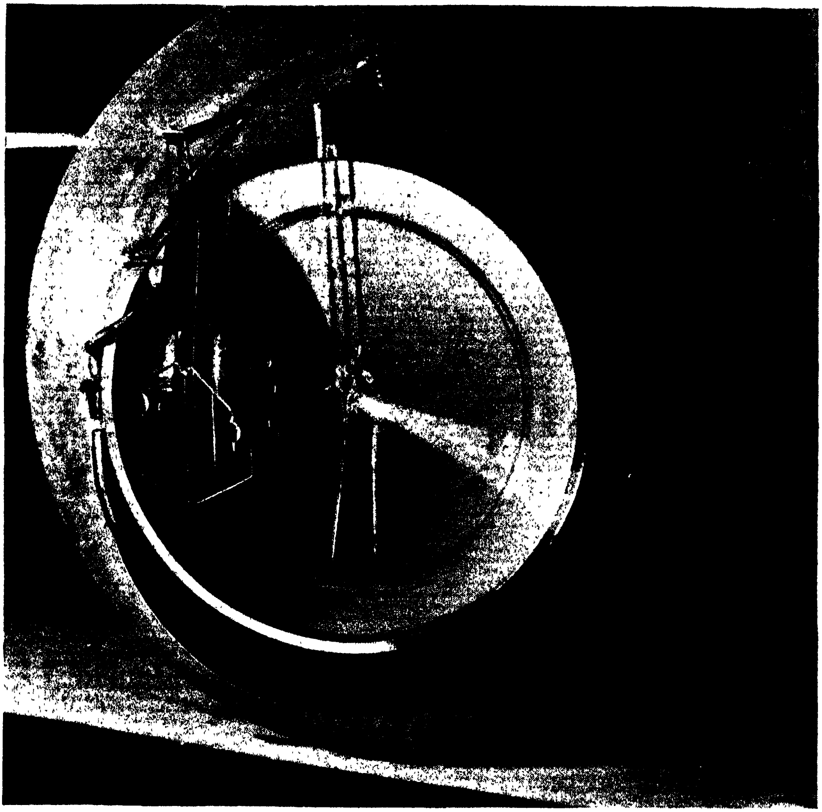


Fig. 64. Centrifugal rotor, pick up, and pendulum.

BEHAVIOR OF THE OIL FILM

We may now ask: "How and with what habit does the oil pass over this rotating surface?" As viewed by the observer, the oil is in a colorless, uniform layer. Even if a brightly dyed oil is passed over the surface, the layer appears almost colorless owing to its thinness. The thickness of the distilland at varying distances from the center has been investigated by two methods. In the first, a deeply dyed oil was passed across the surface, and measurements of reflection density were made. The densitometer was a simple comparator for matching the reflection density of the still surface against that of a standard dilute solution of the dye held in a cell the thickness of which could be varied at will. Figure 65 shows a diagram of the arrangement. Light from a single lamp was allowed to fall in equal intensity on the rotor and on the comparison cell. The rotor, which had previously

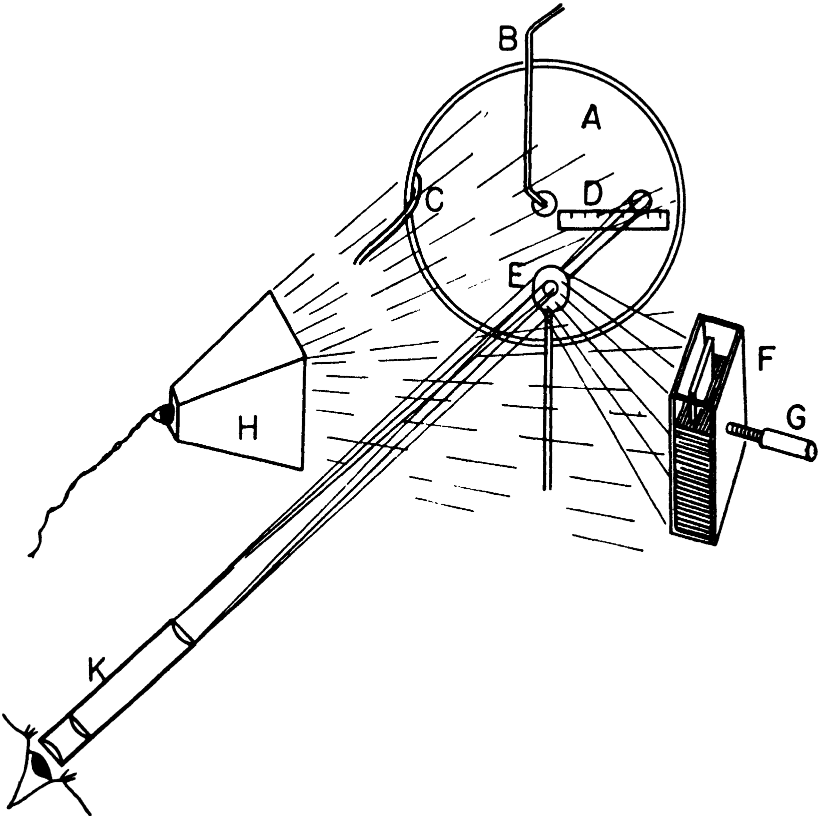


Fig. 65. Density comparator.

been sandblasted to provide a bright matte surface, was viewed through a prism presenting a small central view of any chosen area of the plate. Surrounding this spot the observer saw the reflection from a mirror which was immersed in the cell of dilute dye. The thickness of the dye layer between the mirror and the glass window of the cell was varied by handscrew so that the color seen in the central spot and the surrounding region could be brought to match. In this way the light being reflected from the rotor passed through exactly the same amount of dye as the light coming from the comparison cell. Knowing the relation between the concentration of dye in the oil on the rotor and the oil in the cell, an exact computation of thickness could be made. Typical times of progression across the plate, derived from measurements of thickness, are shown in Figure 66, together with an explanatory legend. The method was explored only for rather low temperatures because all the highly soluble dyes proved slightly volatile under the vacuum conditions and would cloud and color the bell jar during photometry.

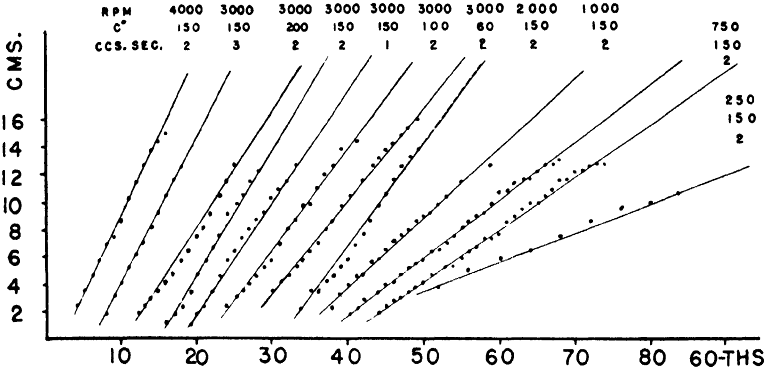


FIG. 66. Rate of travel of distilland. The data are for a plate 36 cm. in diameter and show the time of progression in sixtieths of a second of dyed cottonseed oil over each centimeter of travel from center to edge at different temperatures, rates of feed, and speeds of rotation.

Measurements were continued by the method of the "spiral stream." In Figure 67 you will see that there are two tubes which lead oil onto the spinning rotor. The lower and larger tube carries about nine-tenths of the colorless corn oil employed. The smaller tube applies the other tenth which has been dyed to the limit of solubility with a very opaque blue dye. As both streams reach the plate they begin to proceed outward, but since the dyed oil is applied eccentrically it forms a small separate stream within the main stream, and then the composite pattern begins to spread outward in a regular series of light and dark spirals. From this simple experiment we can learn a number of things of interest to the problem in hand. First of all, the average rate of flow can be determined by counting the number of spirals and dividing this into the revolutions per minute. In the particular experiment illustrated there are 17 spirals and the rotation is 500 r.p.m. The time of travel of the pattern, therefore, is

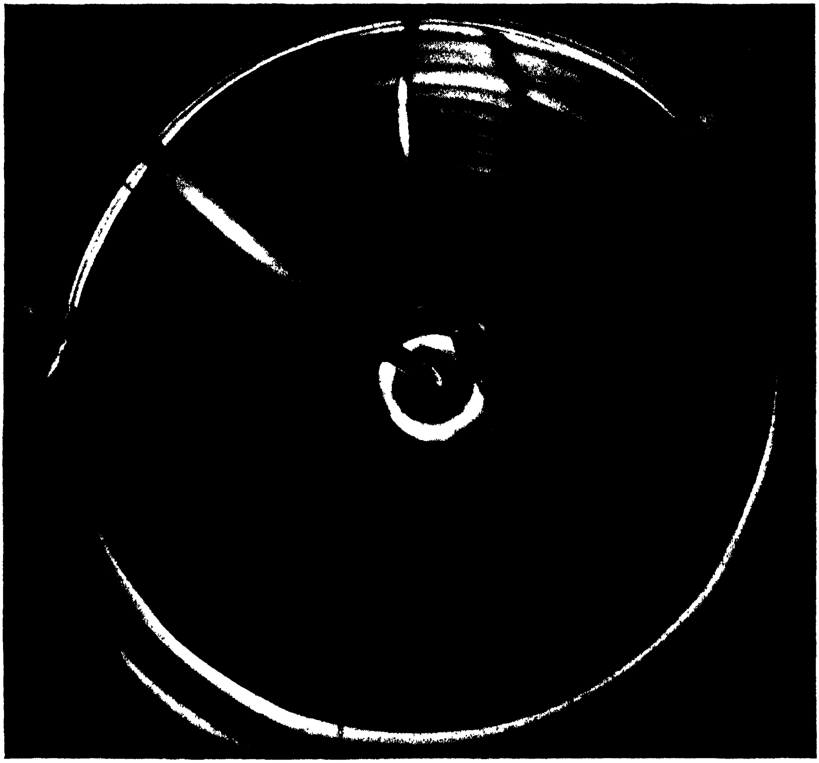


FIG. 67. Photograph of spiral of dyed and undyed oil.

⅓0th of a minute, or two seconds. The plate was 14 inches in diameter and the experiment was done at 100° C. At 250° C. and 3,000 r.p.m., used in actual distillation, the travel is much faster. The photograph also indicates the variation in thickness across the plate. The thickness at any position is directly proportional to the distance between the spirals and inversely proportional to the area of the zone involved. Since the distance is approximately constant throughout we can see that the thickness must vary directly with the distance from the center. This is an unexpected simplification, but it tallies fairly well with the evidence from the reflection measurements. The surprising thing about the photograph is that the spiral pattern persists at all.

According to casual interpretation of classical theory, the pattern should become smudged a few turns from the center. For consider: flow must be either laminar or turbulent. Either the outgoing oil will proceed in irregular tumult or it will proceed in layers with the oil near the plate traveling very slowly, the oil nearing the outside surface traveling fastest, and the oil in between these two regions progressing at a succession of intermediate rates. It will be noticed that the dyed oil has been applied to the bare surface of the plate near the center and is not skidding along the top of the stream of colorless oil. Both streams, therefore, have the same status with regard to the plate. We should expect the blue dye to proceed outward in a succession of layers moving at different speeds, and two or three spiral distances from the center the color should be merged to a uniform blue. That this does not happen merits some study. We interpret it to mean that in the thicknesses involved the rate of diffusion of the dye between one lamination and the next is so rapid compared with the rate at which the laminae slide past that all the molecules have an opportunity of wandering through all the laminations many times on their journey from the center to the edge of the plate. Put in another way, it means that all molecules have an opportunity of coming to the surface many times and being asked that all-important question: "Do you wish to distill?" Phrased in yet another way, our pattern shows us that we have met the critical requirement of spreading our distilland so thin that the evaporating surface shall at any moment represent truthfully the composition of the underlying liquid. Here for the time being we must let the problem rest.*

RADIATION LOSSES

The rotating evaporator has a characteristic which was at first rather disturbing—the rate of evaporation increases with the rate of rotation. Thus with a given heat input, speeding the rotor from 100 r.p.m. to 3,000 r.p.m. quadrupled the amount of distillate collecting on the condenser. It was supposed that friction was the cause or that the metallic rotor was reacting with the electrical heating coil behind. But careful measurements of the input of the motor driving the plate and the input of the electrical heaters showed no discernible change in either. The real reason proved to be the changing radiative properties of the oil in thin layers. It has been stated [10] that the extravagant loss of heat in the molecular still could be prevented by placing polished metal reflectors in front of the evaporator. These expectations have not been realized and the reason is that the polished evaporator and condenser are not the prime radiators and reflectors. The oil itself is the radiator for it is almost completely opaque (we have not yet measured its extinction coefficient) to radiation of 10 to 15 millimicrons corresponding with a temperature of 200 to 300° C. If we could view the oil with eyes sensitive to infrared light of this wave length, we should see, as in Figure 68, Plate I, pitch-black oil becoming "infrared-hot" as it flowed down the evaporator in a falling-film still, for instance. The condensate received by the cooled condenser would also appear black, and as the thickness increased we should fail to see the red-hot oil on the evaporator through the black film of condensed oil. It is evident, therefore, that the wastage of radiant heat is an inherent property of the distilland rather than of the

---

* The considerations given above apply only to small apparatus suitable for laboratory use. The laws governing large volumes of liquid or large rotors are quite different.

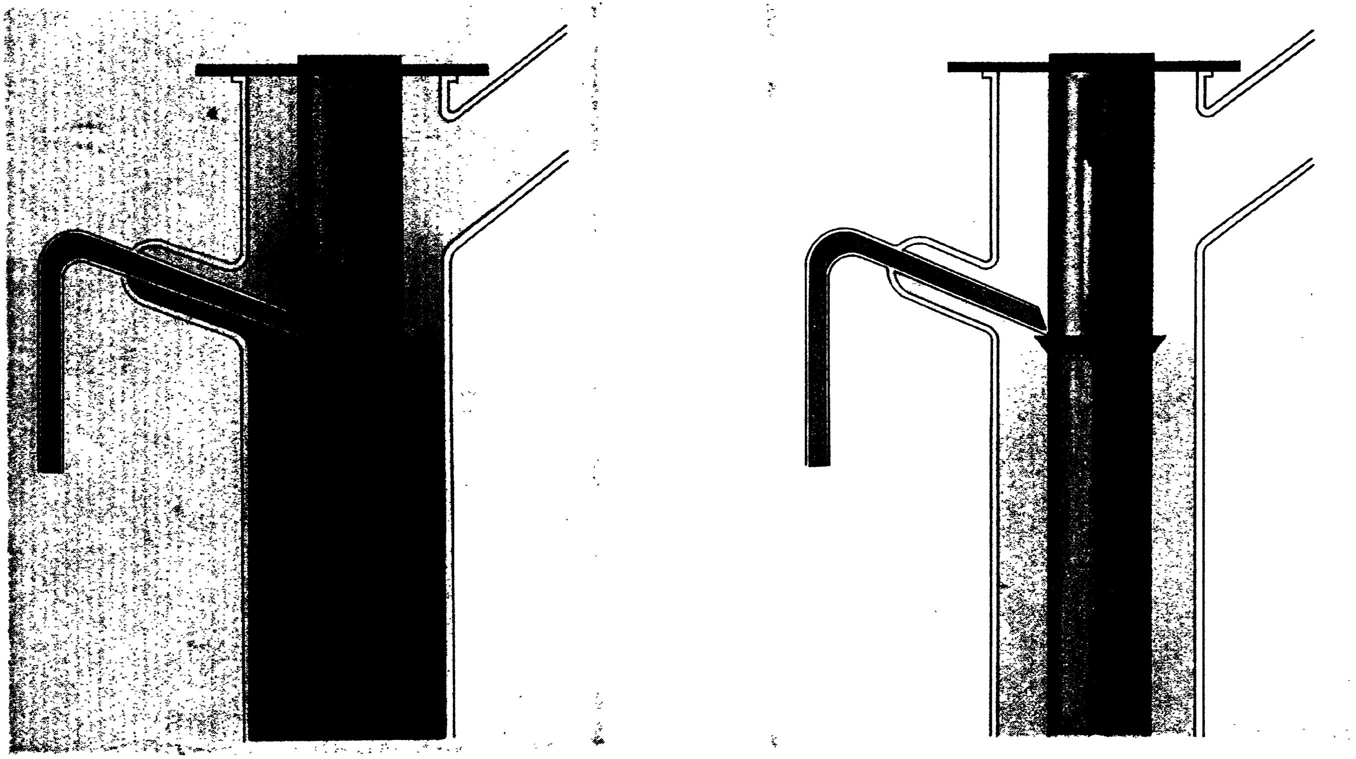


PLATE I—FIG. 7. Artist's conception of falling film still in action as seen by infrared light, left; and as seen by white light, right.

still. When, however, the oil is spread sufficiently thin either on evaporator or condenser, or both, it is possible to minimize loss of radiant heat and to reflect it if so lost. A simple qualitative experiment, illustrated in Figure 69, demonstrates the saving. An electric soldering iron with a thermocouple welded to the underside was mounted on a stand. The iron was maintained with an underside temperature of approximately 200° C. by means of the rheostat, and the iron was periodically wetted by painting it with oil. A sensitive thermocouple was mounted in a similar manner to one side and connected with a galvanometer. Between the iron and the thermocouple a heat insulator was placed. Below the insulator a piece of highly polished nickel was mounted on the shaft of a small high-speed motor. The entire assembly was placed at the edge of a bench so that when the nickel was removed there was a space of some four feet between the assembly and the floor. There were two limiting conditions: one, when the plate was removed and substantially no reflected radiation was received from the iron by the thermocouple; the other, when the brightly polished nickel was placed in position and provided maximum reflection. Intermediate between these two positions were the galvanometer deflections secured when the plate was flooded by various amounts of oil. Depths over 1 mm. all gave about 10 per cent deflection. This evidently corresponded with specular reflection from the top of the oil. With layers less than a millimeter thick the deflection of the galvanometer gradually rose. It was about

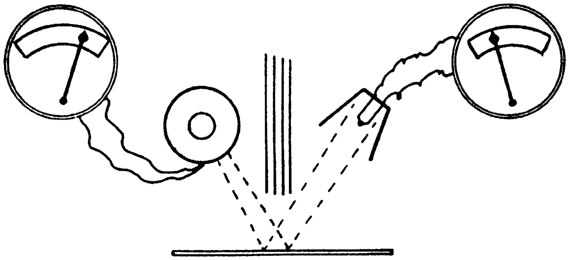


FIG. 69. Experiment demonstrating infrared absorption and reflection by an oil layer.

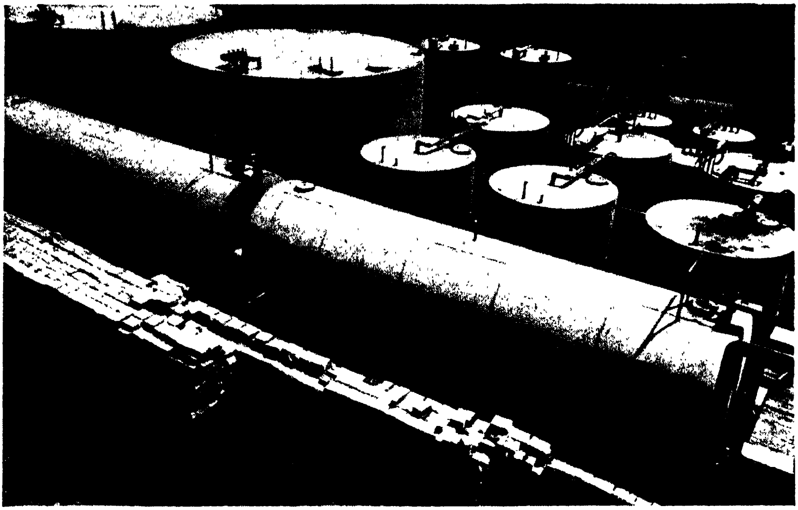


FIG. 70. Fish-liver oil tank farm.

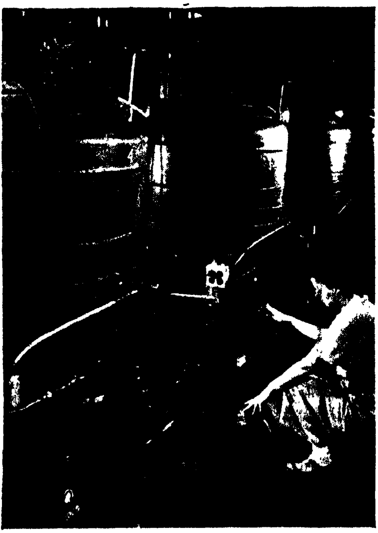


FIG. 71. Some commercial centrifugal stills.

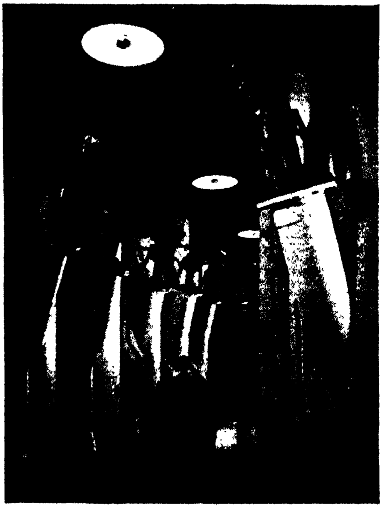
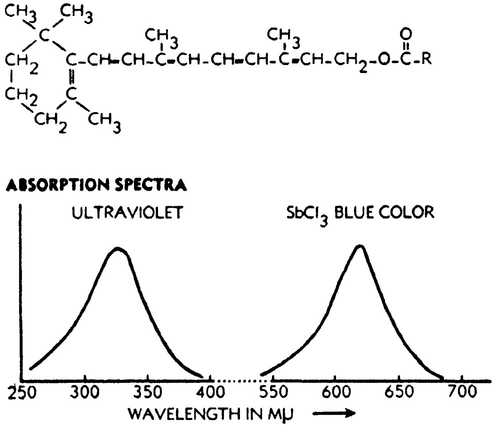


FIG. 72. Blending and storage tanks for vitamin A concentrates.

50 per cent after the nickel plate had been drained on its edge for half a minute. When the plate had been wiped with the finger the reflection increased to 75 per cent of optimum; and only when the plate had been whirled for ten minutes at high speed and appeared entirely free from oil did the reflection approximate that of the polished nickel.

A GLANCE AT THE VITAMIN A INDUSTRY

So much for the evolution of the method. It will have been surmised that the falling-film stills, and later the centrifugal models, have been used in the commercial production of vitamin A. Let us take a look at some of the steps in the process. The first steps are the procurement, testing, blending, and storage of suitable fish-liver oils. The molecular still is an inherently greedy animal as you will see by the size of the tank farm (Fig. 70) needed to satisfy a few stills with rotors 32 inches in diameter (Fig. 71). After the oil has been concentrated it passes to stainless steel blending tanks (Fig. 72) where it is vigorously stirred under inert gas. These tanks toward the end of the summer accumulate a large quantity of distillate containing some tons of vitamin A with a monetary value which may be left to the calculation of the reader.



PROPERTIES:

ELIMINATION MAXIMUM 180°-220° (FOR NATURAL ESTERS)

ADSORPTION ON Al_2O_3 WEAK

$\frac{\text{PET. ETHER}}{\text{83\% ETOH}}$ DISTR. RATIO 4:1 (ACETATE) 99.5:1 (PALMITATE)

CYCLIZES VERY SLOWLY IN ANHYDROUS HCL, GIVING ABSORPTION BANDS AT 352, 369, AND 389 Mμ

BIOLOGICAL ACTION - CURES VITAMIN A DEFICIENCIES

FIG. 73. Vitamin A, formula, absorption spectra, and properties.

Of course, the vitamin A content of the blends must be measured quite accurately. The formula for vitamin A is shown in Figure 73, with absorption spectra. Some beautiful crystals of the vitamin (alcohol), prepared by my colleagues, Dr. Baxter and Mr. Robeson, are shown in Plate II, Figure 74. It will be seen that the vitamin absorbs with a clear-cut maximum at about 330 millimicrons in the ultraviolet. If the vitamin is made into dilute solution in chloroform and rapidly mixed with a concentrated solution of antimony trichloride in chloroform, a brilliant blue color develops for a few seconds which may be measured in the visual or photoelectric colorimeter. The blue color has a sharp peak in the visible at approximately 620 millimicrons. It is therefore customary to measure the concentration of vitamin A concentrates by the use of one of these spectral means. Figure 75 shows our spectroscopist plotting an ultraviolet absorption curve of vitamin A. When she has finished she will measure the height of the curve with a millimeter rule, and the millimeters, multiplied by the "conversion factor" of 2,000 and multiplied again by the price of the vitamin and the weight of the consignment, will determine the size of the bill that the customer will receive. It is evident that quite large sums of money hang on the accuracy of that magnifying glass and the millimeter rule. The customer on receiving the concentrate will probably take an ultraviolet photograph and make his own millimeter measurement. If the two measures agree the money will change hands; if they do not, a prolonged wrangle may ensue. Now there are certain very good reasons for the wrangle. My colleagues have listed a number of substances all of which give the antimony trichloride blue color identical with the one given by vitamin A. Table III shows a list of seven of these and there are still more unlisted. A similar list can be compiled for the ultraviolet absorption measurement. In Table III, only vitamin A and its esters are biologically potent. Anhydro vitamin A has a trace of potency, the other substances probably no potency at all. Obviously, if an oil holds much of these mate-

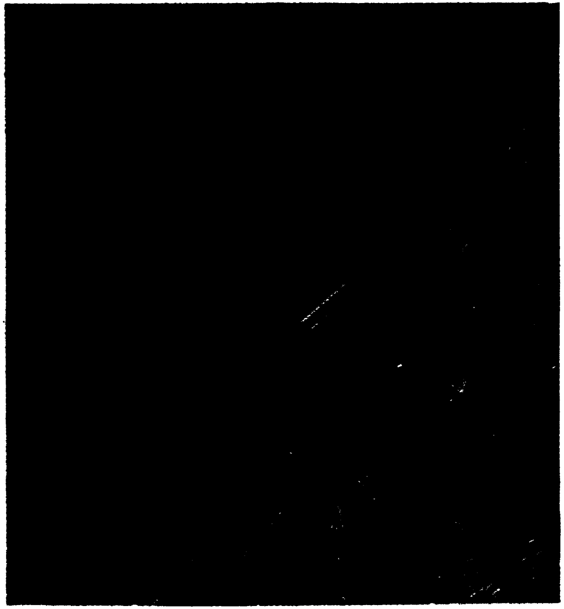


Plate II—Fig. 12. Pure crystalline vitamin A alcohol.

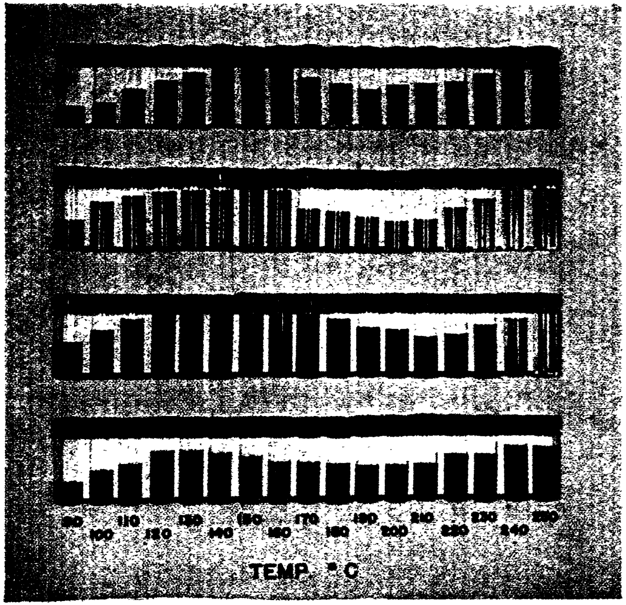


Plate II—Fig. 17. Distillation of pilot dyes. Top Row—Celanthrene red in mineral constant-yield oil. Second Row—Celanthrene red and dibutylaminoanthraquinone in mineral constant-yield oil. Third Row—Celanthrene red and diamylaminoanthraquinone. Fourth Row—Celanthrene red and quinizarin green.

rials the customer will not receive as much biologically active substance as he thinks he is paying for. It is therefore advisable to check the spectrographic assays frequently with biological measurements, as shown in Figure 76.

FIG. 75. Plotting ultraviolet absorption curve of vitamin A in commercial concentrates.

TABLE III

Substances which Absorb Utraviolet Light and whose $SbCl_3$ Blue Colors have Absorption Maxima at 620 Millimicrons

Vitamin A
Vitamin A Esters
Anhydro vitamin A
Iso-anhydro vitamin A
Pseudo vitamin A
Sub-vitamin A
Anhydro sub-vitamin A

RESEARCH WITH THE MOLECULAR STILL

As remarked earlier, the commercial production of vitamins demands collateral research into many fields which are at least as intriguing as the parent process. These researches form the second part of this article. Before describing them, however, it is necessary to explain the technique of the *elimination curve* [11] by which quantitative molecular distillations are done in laboratory apparatus.

Distillation is done to purify, to separate mixtures. As one constituent after another appears in the distillate, it is necessary to change the receivers. Distillation must do two things, therefore. It must provide both separation and information as to when to change the receivers. This information is obtained in ordinary distillation by observations of boiling points. A "point" is defined by the crossing of two lines or values. The two values in ordinary distillation are pressure of vapor, which is assumed to be equal to the pressure of the residual gas, and temperature; the *thermometer* and *manometer* are the two interlocking instruments. In molecular distillation there is by definition no residual gas pressure. Molecules are leaving the liquid at some rate at all temperatures. The rate is generally computed by Langmuir-Knudsen equation:

FIG. 76. Bioassay of vitamin A.

$$n = PA\sqrt{\frac{1}{2\pi MRT}}, \text{ where}$$

n = rate in moles per second,
P = vapor pressure, dynes/sq. cm.,
A = area of distilling surface, sq. cm.,
M = molecular weight,
R = ideal gas constant, and
T = temperature, °K

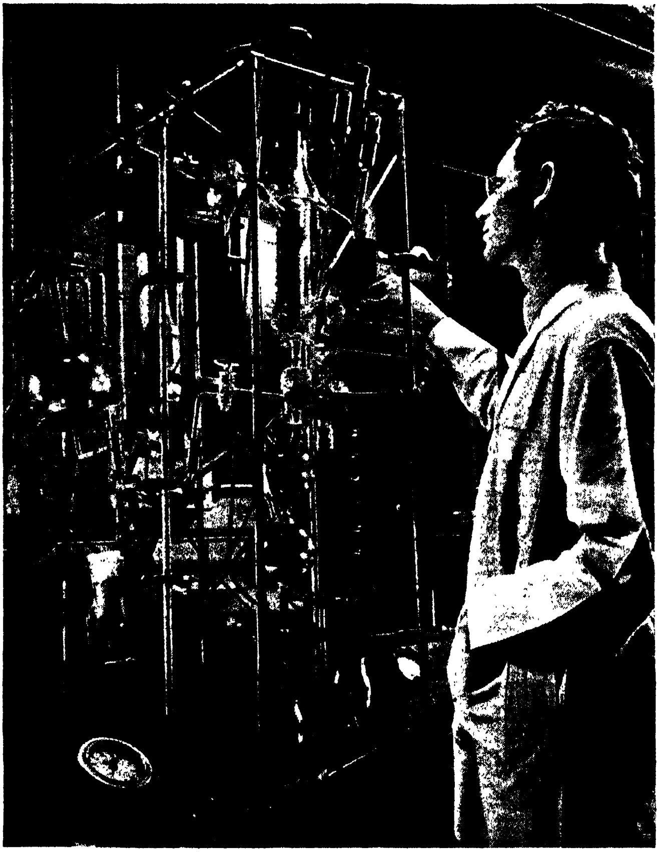


FIG. 77. Cyclic batch still for quantitative molecular distillation.

It will be seen here that the number n of molecules evaporating at any moment is proportional to the measured vapor pressure P, multiplied by area over which they are evaporating, into the square root of the absolute temperature, the weight of the molecules, and the gas constant. We must therefore measure, as Burch pointed out long ago, rate instead of pressure, and the two interlocking instruments for telling when to change the receiver are the *thermometer* and the *clock*. Quantitative molecular distillation apparatus therefore involves the accurate use of time and temperature. A photograph of the apparatus with operator and clock is shown in Figure 77. Up above there is a falling-film evaporator. This, of course, can be replaced by a centrifugal evaporator or any other unit evaporation surface. Below are two reservoirs for distilland and a metering pump for circulating the distilland from one reservoir over the evaporator to the other reservoir. There are means for changing the receivers without breaking the vacuum. This apparatus is quite tedious to use since when once distillation has been started it is necessary to take all the fractions off without stopping the procedure, otherwise very serious drainage errors are introduced.

The rules for cyclic, quantitative distillation are simplicity itself:

1. The distilland must be in *liquid* form to insure proper selection of distilling molecules.
2. The volume of distilland should not diminish to less than half during the entire course of distillation.
3. The volume of distillate coming off at each cycle should be approximately the same. This obviates serious drainage errors.

Since not all materials fulfill these requirements, it is often necessary to add to the material being distilled quantities of *residue** and synthetic *controlled-yield-oils* to adjust the rate of distillation. The active material under study can be extracted from the distillate oils afterward.

4. Temperature must be advanced by a regular scheme. This may be conveniently equal temperature intervals or geometrical increments of absolute temperature. We have found equal temperature intervals quite satisfactory.
5. The time of cycling must be maintained constant.

---

* Relatively nonvolatile oils which have previously been stripped of volatile constituents.

In summary, therefore, a specially prepared solution of the product under distillation is cycled for equal intervals of time with equal increments of temperature and a large number of fractions are taken off until it is certain that everything has been distilled.

As an exercise in cyclic distillation, we may take any pure substance within the distillation range. A substance convenient to our laboratory is vitamin A. Pure crystalline vitamin A, to make say a tenth of a per cent solution, is blended with constant-yield and residue oils and introduced into the still. The distillation is begun at 70° C. and fractions are taken off at 10° intervals until a temperature of about 210° C. has been reached. The fractions are then examined for vitamin A, as by the antimony trichloride test, and the potency of vitamin A is plotted against temperature or fraction number, as shown in Figure 78. It is found, rather disappointingly, that there is *some* vitamin A in every fraction between 70 and 210; that the concentration of vitamin A rises to a pronounced maximum at a fraction near the center.

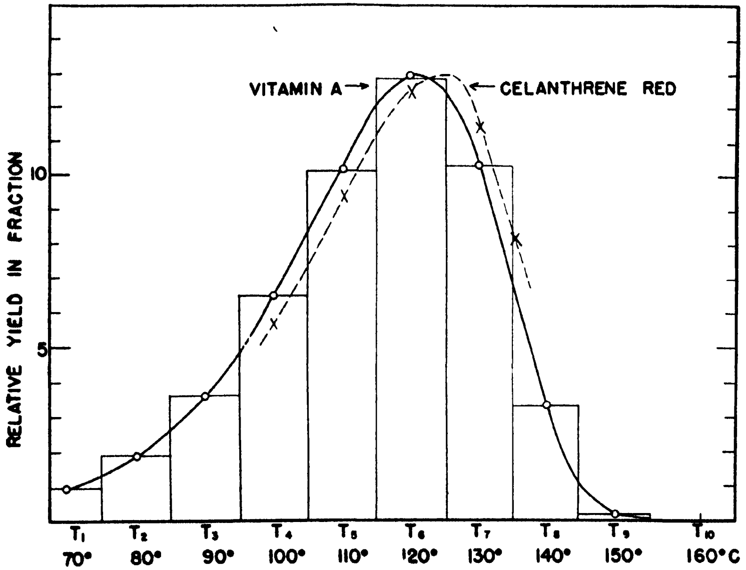


FIG. 78. Elimination curve of vitamin A.

It is evident that it is not feasible to separate two substances completely in one distillation unless they boil more than 100° apart. This gives us a graphic idea of the limitations of molecular distillation. The still is valuable because it does distillations at very low temperatures, not because it gives good separations; it gives poor separations.

THE ELIMINATION CURVE

We shall now examine why the potency of vitamin A varies in this particular manner. Let us adjust the still so that for the first distillation cycle one part in a hundred of the vitamin A evaporates. This we shall call a distillation rate of unity. Now let us raise the temperature 8 to 10° C., which will be approximately the increment needed to double the rate of distillation. Next time we shall expect to find in the receiver two parts of vitamin A, or rather the amount of vitamin A that will evaporate with double the rate of distillation applied to the 99 parts left in the still. Next time the temperature is raised another 8° and now the intrinsic distillation rate is 4, but that "4" operates on approximately 97 parts left in the still. Next time the intrinsic rate will be 8, operating on about 93 parts. Next time it will be 16, operating on less than 80 parts and so on. Eventually the intrinsic rate of distillation becomes indefinitely high but the distilland has been depleted of so much of its vitamin A that the product of rate and concentration approaches zero. The concentration of vitamin A in the distillates thus passes through a maximum according to a well-defined progression or *elimination curve* [12] which is dependent not on the chemical nature of vitamin A but on the *history of the events of distillation.* It can be assumed with some certainty that all the materials placed in the still will distill with exactly this kind of curve, providing, of course, that there are no complicating factors such as temporary combination with the distilland and later liberation at higher temperatures. With this general rule established it is evident that a number of substances in a mixture will all come to maxima at temperatures typical of their rate of evaporation. We shall have a number of overlapping elimination curves, such as shown in Figure 79 for the di-substituted aminoanthraquinone dyes, for instance. This series also shows that we have come into the evaporation range in the molecular still of the brilliantly colored aniline dyes. Dyes can therefore be used as pilots for distillation. If a few parts per million of a colored dye are placed in the oil this will come out in its characteristic place in the distillation series. It is possible to match the invisible substance in which we are interested with a visible pilot. Plate II, Figure 80, shows what beautiful gradations can be secured by distilling the appropriate dyes. In the top row hydroxyaminoanthraquinone (celanthrene red) is distilled from mineral constant-yield and residue oils. Row two shows dibutylaminoanthraquinone mixed with celanthrene red, row three diamylaminoanthraquinone and the red dye, while the bottom row shows a complete separation of celanthrene red and quinizarin green. Note that each dye comes to its characteristic maximum independently of the others and of

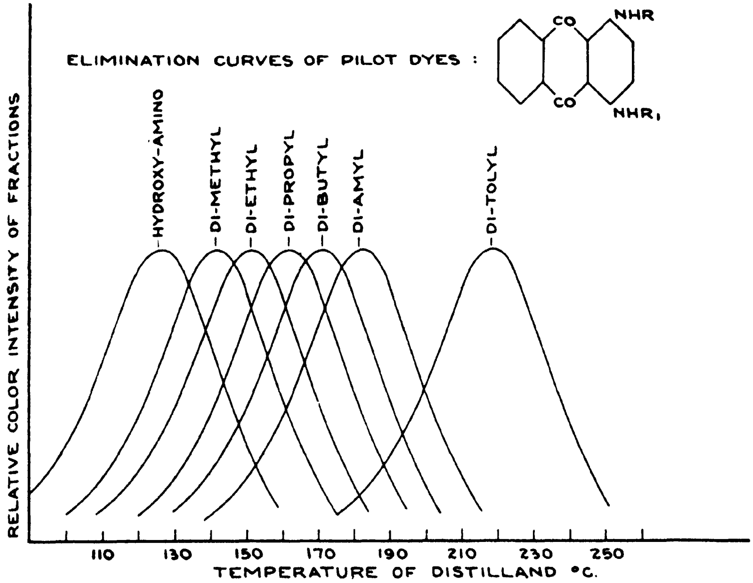


FIG. 79. Elimination curve of homologous series of anthraquinone pilot dyes.

the rate of distillation of the base oil. Colored constituents often separate out in serial order on the condenser in commercial distillations. Thus we may find a crystalline deposit of sterols colored in concentric rings with the yellow and orange of various marine carotenoid-like substances. An experimental distillation done with vegetable sterols, celanthrene red and quinizarin green, all dissolved in corn oil, is shown in Plate III, Figure 80a.

Using the elimination method on vitamin A, the vitamin

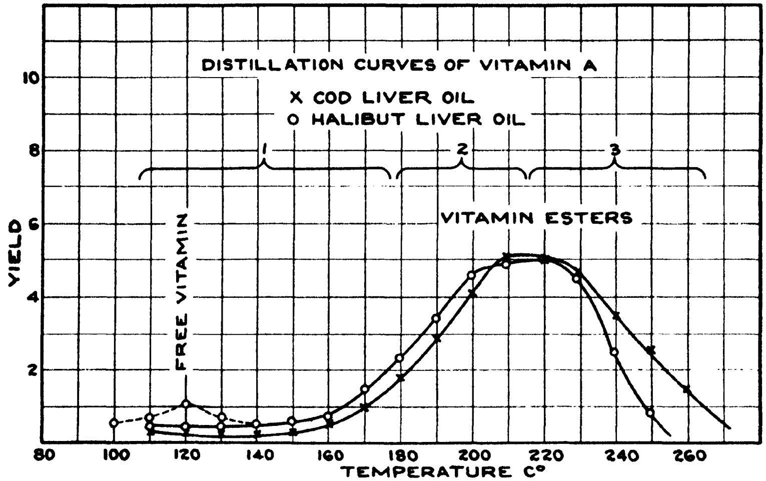


FIG. 80c. Distillation curves of vitamin A fish-liver oils.

and celanthrene red are dissolved in a mixture of residue and constant-yield oils and distilled. The two substances appear in the distillates in maximum yield at about 127° C. Having established that vitamin A has a maximum at 127, we can then investigate such oils as cod liver and halibut liver in which we are interested. Celanthrene red is placed in these oils and a curve is obtained which, however, completely negatives the postulates we have just deduced. Instead of a maximum of 120 we find one spreading broadly between 200 and 230 which is obviously of the wrong shape (Fig. 80c). A little investigation shows that we are dealing here with bound vitamin A. The vitamin occurs in natural oils esterified

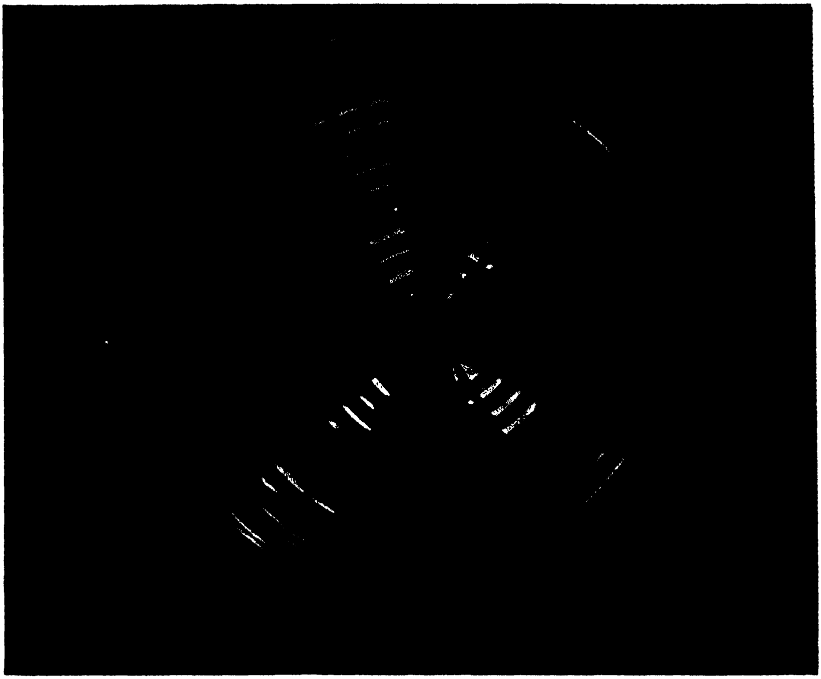


PLATE III—FIG. 18. Separation of pilot dyes (celanthrene red and quinizarin green) on experimental condenser placed in front of centrifugal plate in large cyclic batch still.

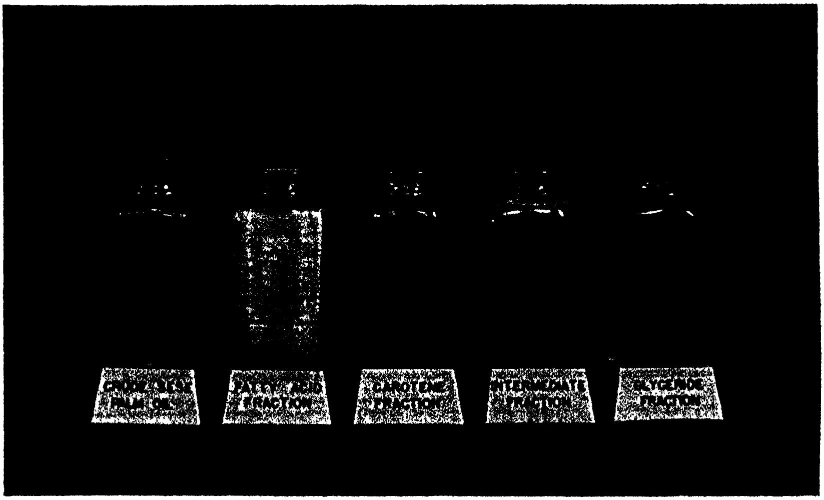


PLATE III—FIG. 19. Samples from a commercial distillation of crude palm oil. The fatty acids distill at the lowest temperature and are followed by a rich carotene fraction. More of the carotene distills in the next fraction with the solid fats, leaving a final pale yellow glycerine fraction which is liquid at room temperature.

with acids representative of all the fats that are in the liver oil. There is a small maximum at about 120 which is evidently due to anhydro vitamin A. When these natural oils are saponified with alkali and the ether extract transferred to plain constant-yield and residue oils and redistilled then we find a new maximum at the correct place and with the correct curve.

This method has been applied to vitamin D with the rather curiously shaped curve shown in Figure 81. Here it will be noticed that there are three approximately parallel curves. The reason for the three is that no convenient test is known which is sensitive to vitamin D in very high dilution. The points on the curve were therefore determined by rat assay; that is to say, by feeding small quantities of the distillate to rats depleted of vitamin D, and observing whether rickets is cured. Each point on the curve is thus the graveyard of a score of rats and the entire curve is a cemetery of vast proportions. The central line is the best value deducible from the assays and the two outer lines represent the limits beyond

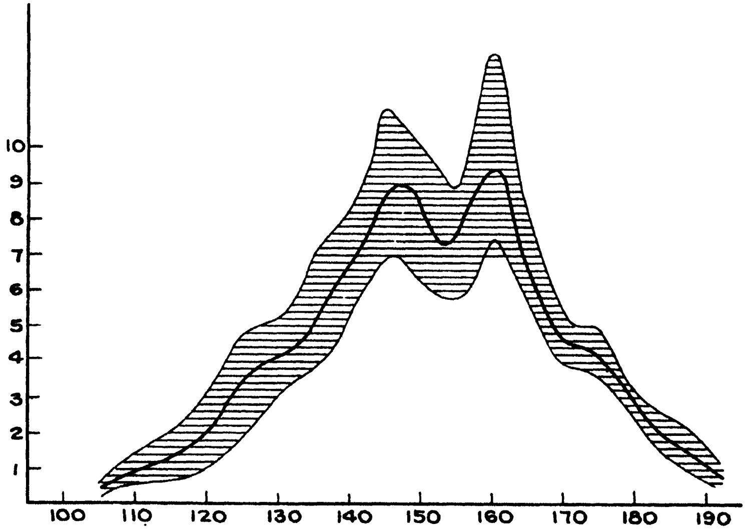


Fig. 81. Master elimination curve of natural vitamin D. The central line represents the lowest probable potencies of the fractions as examined by rat assay. The shaded area represents the probable limits of uncertainty of assay.

which the assay data cannot reasonably be interpreted. Even so, it is scarcely possible to draw a normal elimination curve within the limiting curves and there is considerable evidence of more than one kind of vitamin D in the tuna-liver oil used.

The biochemist examining this curve will say at once that it represents not a number of D vitamins but the errors inherent in animal assays. Fortunately there is an excellent method available for testing this contention. If the alleged low-boiling D vitamins distilling at between 100 and 120 are identical with the alleged high-boiling vitamins distilling between 180 and 190, then a redistillation of these fractions, in miniature, and again bioassaying should yield sub-curves which are replicas of the master curve. Accordingly, two redistillations were performed. Four of the earliest fractions were blended and four of the highest fractions were blended with carrier oil; the blends were redistilled and the fractions assayed. It will be seen from Figure 82 that considerably different curves were obtained. On the low-boiling side the small irregularities of the master curve have become swollen. The same thing has happened to the high-boiling side of the other

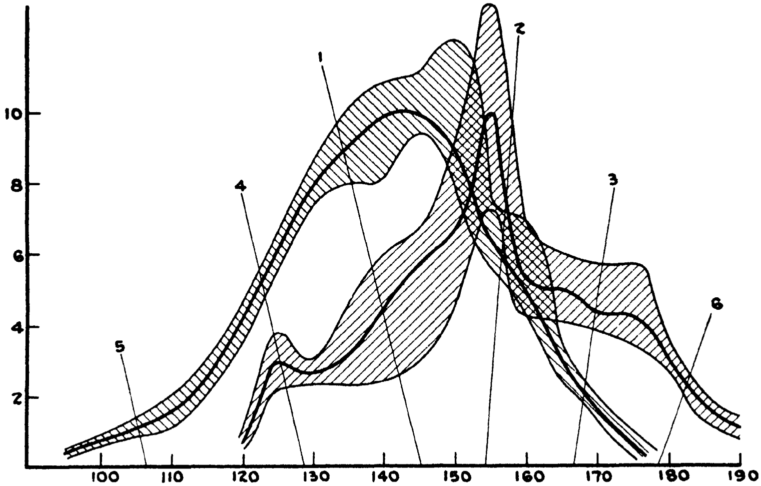


Fig. 82. Redistillation curves of natural vitamin D. The curve to the left is for the low-boiling fractions of the master curve, Fig. 81; the right-hand curve is for the high-boiling fractions.

curve. Later Dr. Bills [13] tested the rat:chicken potency ratio of the lowest boiling distillate and found this to be quite different from the parent cod-liver oil.

It will be remembered that vitamin D_2 is able to cure rickets in mammals but not in birds, whereas vitamin D_3, the marine oil vitamin, is potent both for poultry and rats. The general formula for the D vitamins is shown in Figure 83: where R is an aliphatic chain of varying complexity. It is evident that the lowest boiling vitamin must have a very short chain if the rest of the molecule is to be considered intact; and we wondered if, turning the formula upside down, there might be a fortuitous relationship like Figure 84, but no, the length of the chain bears no relation to the length of the animal's tail!

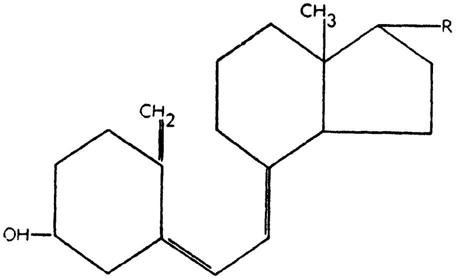


FIG. 83. General formula for the D vitamins.

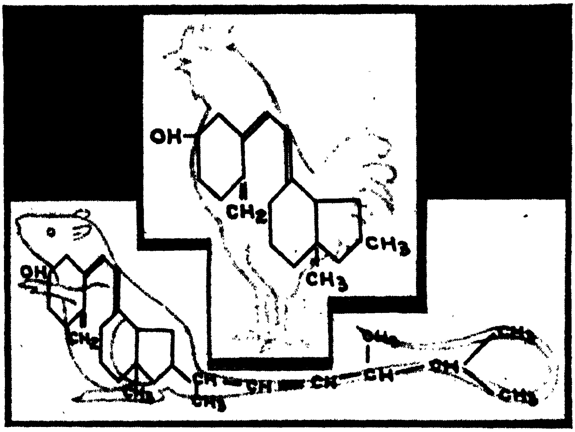


FIG. 84. Hypothetical "rat-chicken" structure of high-boiling and low-boiling vitamin D.

The distillation just described has had no commercial or medicinal significance, since an excellent "natural" vitamin D has been synthesized and is now available. It does, however, pattern a means for determining purity. We may make the broad generalization that an organic chemical, or indeed any substance is to be considered pure only until someone proves it is impure. The method of redistillation of high and low fractions provides a criterion of purity from a distillation point of view. Thus if the reader has prepared a new complex drug and this has been supplied to the medical profession who report excellent results with it, except that they are sometimes contradictory, the question might arise, "Does the drug act in two ways or does it contain more than one substance?" A means of finding out would be to put the drug in a suitable solvent, distill it in the cyclic still and then subject the high and low fractions to redistillation, and also furnish them to the doctors for retest.

SOME STUDIES WITH FISH-LIVER OILS

We have now illustrated the elimination curve method sufficiently to put it to practical application. Figure 85 shows a distillation map of a typical fish-liver oil and depicts the substances of chief human interest. No. 1 is the odor. This distills out at the lowest temperatures and collects largely in the freezing traps of the high-vacuum pumps. The reader would be astounded at the potency and magnitude of the "smell" that collects from tons of marine oil during the course of a year. Next in order, No. 2, are the fatty acids, and the broadness of the curve shows that when a marine oil goes rancid the acids from all the fats hydrolyze simultaneously. No. 3 is the curve for spurious or anhydro vitamin A and similar substances. Curve No. 4 is the cholesterol and esters, vitamin D_3 and its esters; No. 5 shows vitamin A alcohol and esters. No. 6 is the natural preservative substance which after considerable study has turned out to be α-tocopherol, or Vitamin E [14]. One might ask, "Why does a fish have vitamin E in its liver?" "Is it for some reason connected with fertility?" Our guess is "No"; that the vitamin E is present in the liver oil for the same reason that it is present in vegetable oils and in animals and human beings, namely, to prevent rancidity. The shark living immersed in water requires considerably less vitamin E than the rest of us who are maintained at 37° and exposed to the atmosphere in the summer. We have often wondered and often been asked why fishes in general have so much vitamin A in their livers, why the soupfin shark, for instance, carries a million times more vitamin A than a human being. The best answer we could think of until recently is that marine animals living as they do in chilly water require the vitamin to keep them from catching colds! A more reasonable but still tentative hypothesis is that vitamin A, being an active pro-oxidant, is required by the fish to stimulate the utilization of oxygen under the lower solution pressure and potential of oxygen existing under water. If this is the case a correlation should be found between vitamin A content and habitat, as well, of course, as between species of fish.

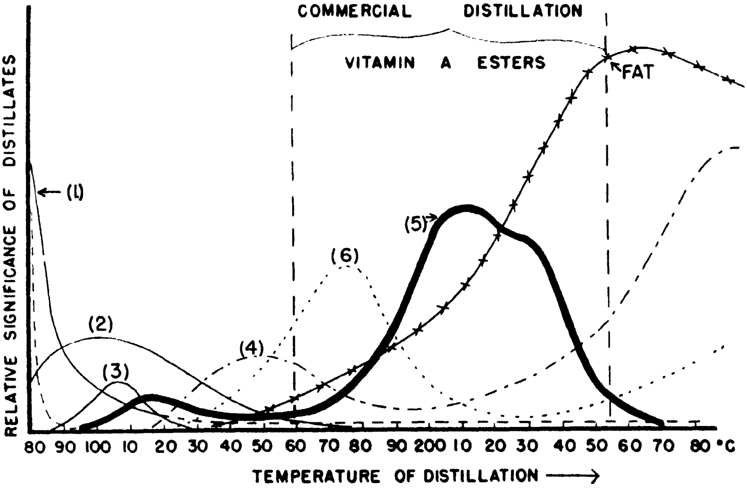


FIG. 85. Distillation map of some constituents of a fish-liver oil.

A few years ago Russian and British chemists [15] discovered vitamin A_2. It was found in the retinas and later in the livers of fresh-water fishes. Why fresh-water fishes require a different material for their visual purple than the marine fishes, and why human beings and animals should follow the marine rather than the fresh-water fishes are still open to speculation. However, the two kinds of vitamin do exist, and they have been shown to be very nearly related

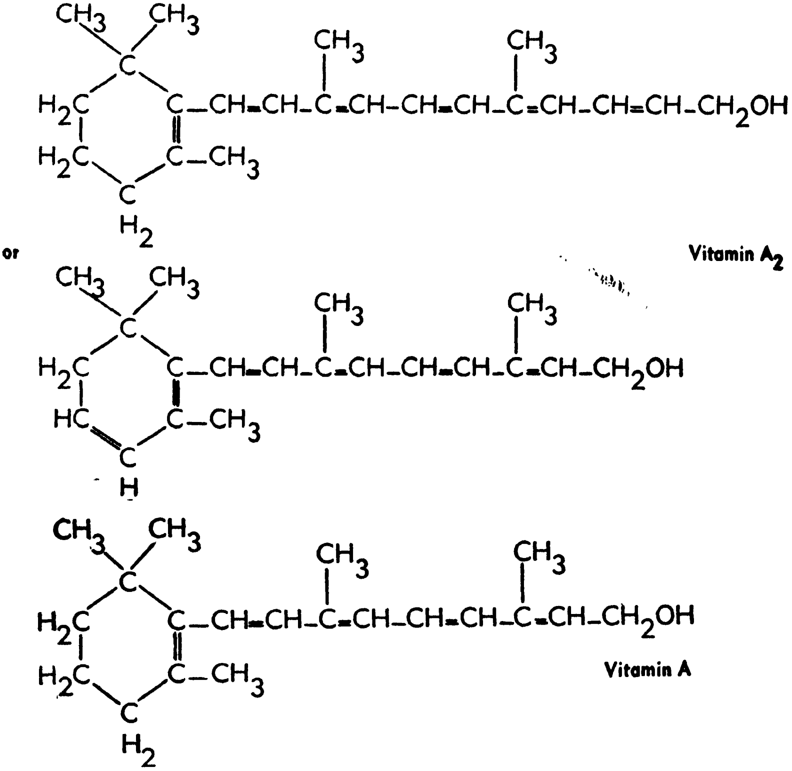


FIG. 86. Probable formulae for vitamins A and A2.

chemically. In fact, their formulae are suspected to be as shown in Figure 86. It is a question which, if either, of the alternative formulae proposed for vitamin A2 is the correct one. Now there is one important difference between the two: the first one requires two extra carbon atoms to provide the extra double bond postulated. We have previously shown that each carbon atom added to an aliphatic chain in a series of homologous compounds raises the elimination maximum by approximately 4½°. Therefore the methylene group added to vitamin A to form vitamin A_2 [16] should raise the elimination maximum by 9°. Dr. E. LeB. Gray and Mr. J. C. Cawley in our laboratory have determined the elimination maxima of A and A_2 in comparison with celanthrene red and showed that the difference in boiling point was 1½° at most and they might possibly boil at the same temperature. This would appear to dispose of the theory that there may be two extra carbon atoms in the new vitamin.

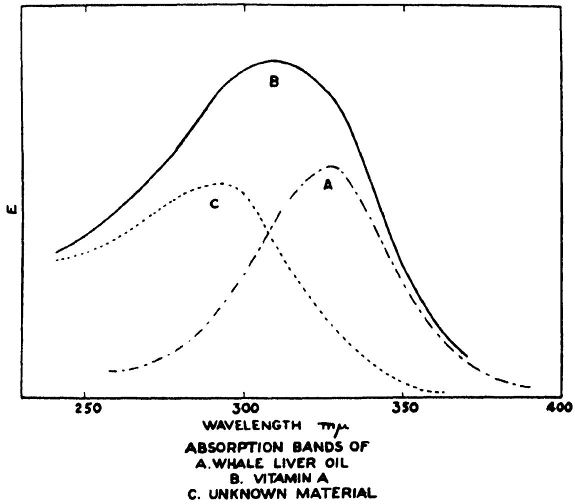


FIG. 87. Absorption spectra of whale-liver oil, vitamin A, and kitol.

By far the most dramatic discovery [17] that has so far come from the molecular still has been arrived at by the process of quantitative molecular distillation just described. For a long time whale-liver oil has been esteemed as a rich source of vitamin A. Indeed, it has been used by the British and Danes for many years for the fortification of margarine. When supplies were withdrawn from Denmark during the last war, a critical shortage of vitamin A developed. Elaborate precautions are being taken during the present war to avert the same conditions. Whale-liver oil has not been favored as a source of vitamin A in this country partly through lack of availability but chiefly because the ultraviolet spectral absorption curve is not that of pure vitamin A. It had been realized for some time that the vitamin A present was probably normal in nature and that there was another material which was distorting the curve. Diagrammatically speaking, the Curve B of Figure 87 would be composed of Curve A for vitamin A and Curve C for an unknown substance. But nobody did much about isolating the unknown substance. Embree and Shantz set about the matter in a characteristic manner: they saponified whale-liver oil, placed it in carrier oil in the molecular still, and then took it over that part of the elimination cycle which should take out all the free vitamin A. The distilland left in the still should therefore have an absorption spectrum of the substance *C*, and this proved to be the case. There was, however, no indication of the strength or potency of the material *C*, or what percentage solution it represented. They therefore decided to distill the substance *C* at a higher temperature, and at about 220° distillation became vigorous. The authors then examined the golden-yellow distillate in the ultraviolet spectograph and to their consternation found that they were dealing with a strong solution of vitamin A without any of the spectral material *C* showing. This was all the more extraordinary, since they knew they had previously distilled the whole of the vitamin A out of the solution. They then examined the twice-spent distilland and found that the spectral absorption of *C* had diminished to about half. They put it back in the still, performed the distillation again, and obtained a new yield of vitamin A in the receivers. After much experimenting it became evident that they were distilling a precursor of vitamin A which decomposed at the temperature of the molecular still, liberating the volatile vitamin A which was thus removed not only from the reaction mixture but from the zone of harmful decomposition, since vitamin A cannot stand a temperature of 250° C. for any prolonged period.

The new substance was named "kitol" (*ketos*, a whale) and launched on its scientific career. Kitol is particularly interesting because it is the mirror-image in properties of β-carotene. Carotene, the normal precursor of vitamin A, is split into the vitamin in the animal body but no one knows how to do it in the laboratory. Kitol, on the other hand, is apparently not split into vitamin A in the animal body but can be turned into the vitamin by high-vacuum decomposition. One may speculate that kitol is not a provitamin but a *post*vitamin. We do not yet know whether the kitol is present to any extent in the fresh whale liver or whether all of it is formed during the period which elapses between catching the whale and rendering the liver. In any event it is evident that the whale, being a mammal instead of a fish, is probably not adapted and does not have the requirements for the enormous quantity of vitamin A which the fishes themselves require or accommodate. The whale is, however, forced to live on marine food, plankton and small fishes, etc. and has the same basic diet as the fishes themselves. Apparently there are dietary problems in a whale's life and one can imagine the whale complaining that it is compelled to eat this wretched sea food again tomorrow and the next day and wondering what to do about it. Perhaps the whale removes some of the excess vitamin A as kitol. Embree and Shantz have found an equivalent substance, kitol$_2$, in the livers of fresh-water fishes, and kitol has been found in mammalian livers. Whether the kitol is a safety valve for vitamin A or whether again it represents a final line of defense against vitamin A deficiency is not yet decided.

VITAMIN CHANGES IN THE ANIMAL BODY

The molecular still has aided in tracing the different states of combination which substances of biologic potency acquire in their passage through the animal body [18]. For instance, most of the vitamins, sterols, and hormones pass interchangeably to free and ester forms, the latter making themselves known by distilling at much higher temperatures.

Changes in the body to vitamin A have been investigated in this manner. The vitamin is normally ingested through food or medicine in many different forms. There are the α- and β-carotenes and cryptoxanthine of plants. There is free vitamin A alcohol, the vitamin A esters such as found in marine-liver oils, and there are the vitamin A acetate, palmitate and succinate of the synthetic laboratory. All these substances if distilled would show elimination maxima at different places. Question: When they are swallowed by rat or man, do they all appear as the same kind of vitamin in the body and are they stored as the same vitamin in the liver, or is each one stored and utilized in the form given? This matter was tested by feeding six variations of vitamin A to six groups of rats. After a considerable dose of the vitamins, the animals were killed, the livers extracted, the extract placed in constant-yield oil, distilled, and the $SbCl_3$ blue color plotted against fraction number, as in Figure 88. It will be noticed that all the different kinds of vitamin A were converted into substantially the same kind of vitamin A mixture in the livers. Thus each group contained a small quantity of

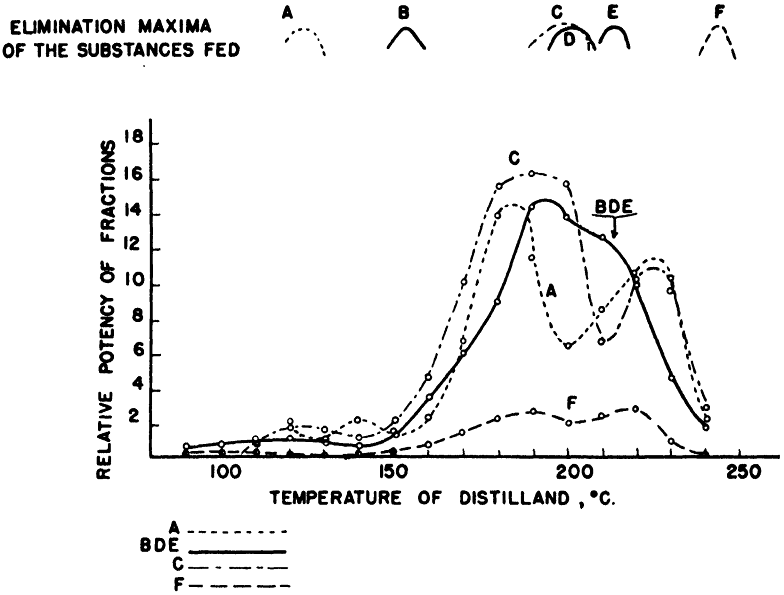


FIG. 88. Distillation curves of vitamin A preparations recovered from the liver of the rat.

anhydro and free vitamin A, a large quantity of vitamin A-palmitate-stearate and a small but still pronounced quantity of a C_{22} ester of vitamin A. Even the β-carotene which had been absorbed in only limited amounts reappeared as the same two kinds of high-boiling esters with a small quantity of free vitamin A. The idea, therefore, that the C_{22} acid comes from the marine oil itself is unfounded because β-carotene provides exactly the same kind of ester. We have not had occasion to perform these distillations on vitamin A recovered from many animals and have not had the privilege of distilling the vitamin from human liver as yet, but there is little doubt that each species makes its own characteristic ester map. We look forward to the day when the sleuth in the scientific detective novel will make his deduction as to the identity of the murdered man by extracting the liver, placing it in constant-yield oil, distilling out the vitamin A esters, examining the curve with his magnifying glass, and trimphantly announcing that the victim was a Chinese coolie.

The distillation of vitamin A stored in the liver is an extremely interesting problem, the investigation of which is by no means complete. A few years ago Moore in Cambridge, England [19], showed that vitamin E ingested simultaneously with vitamin A increased the storage of vitamin A in the liver. Harris and Cawley in our laboratory have distilled vitamin A recovered from the livers of two groups of rats, one of which was given vitamins A and E while the other received vitamin A alone. It is significant that when vitamin A is given alone a C_{18} type of vitamin A ester is recovered from the liver. When vitamin E is administered with the vitamin A, the C_{18} esters are accompanied by larger quantities of C_{22} esters of vitamin A, and the total storage in the liver is increased. Since the C_{22} esters evidently contain highly unsaturated acids, it would appear that the vitamin E has acted as a preservative not only for the vitamin A but for unsaturated fats which may also be in the class of indispensable food factors. Harris, with J. Jensen, is pursuing the matter further.

At the present time our laboratory is occupied with studies of the sparing action of the E vitamins on one another and food or medicine in many different forms. There are the *α*- and *β*-carotenes and cryptoxanthine of plants. There is free vitamin A alcohol, the vitamin A esters such as found in marine-liver oils, and there are the vitamin A acetate, palmitate and succinate of the synthetic laboratory. All these substances if distilled would show elimination maxima at different places. Question: When they are swallowed by rat or man, do they all appear as the same kind of vitamin in the body and are they stored as the same vitamin in the liver, or is each one stored and utilized in the form given? This matter was tested by feeding six variations of vitamin A to six groups of rats. After a considerable dose of the vitamins, the animals were killed, the livers extracted, the extract placed in constant-yield oil, distilled, and the $SbCl_3$ blue color plotted against fraction number, as in Figure 88. It will be noticed that all the different kinds of vitamin A were converted into substantially the same kind of vitamin A mixture in the livers. Thus each group contained a small quantity of

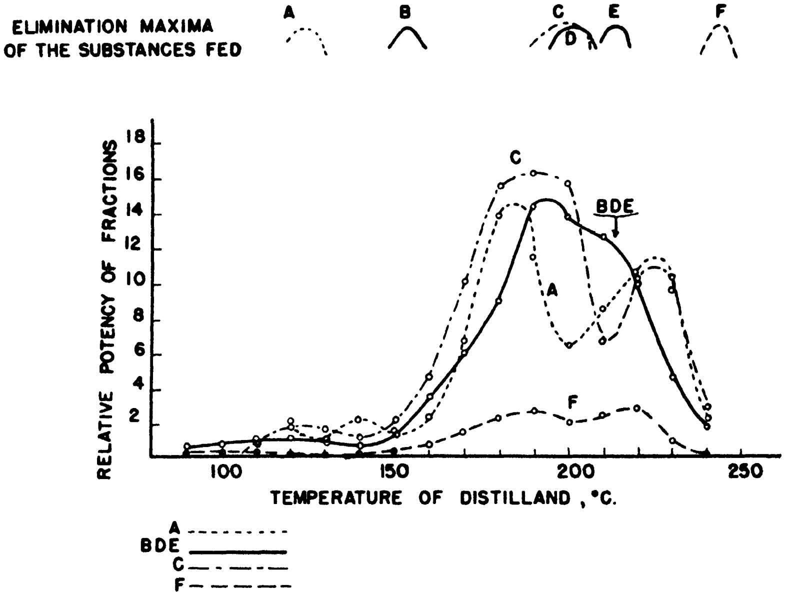


FIG. 88. Distillation curves of vitamin A preparations recovered from the liver of the rat.

anhydro and free vitamin A, a large quantity of vitamin A-palmitate-stearate and a small but still pronounced quantity of a C_{22} ester of vitamin A. Even the β-carotene which had been absorbed in only limited amounts reappeared as the same two kinds of high-boiling esters with a small quantity of free vitamin A. The idea, therefore, that the C_{22} acid comes from the marine oil itself is unfounded because β-carotene provides exactly the same kind of ester. We have not had occasion to perform these distillations on vitamin A recovered from many animals and have not had the privilege of distilling the vitamin from human liver as yet, but there is little doubt that each species makes its own characteristic ester map. We look forward to the day when the sleuth in the scientific detective novel will make his deduction as to the identity of the murdered man by extracting the liver, placing it in constant-yield oil, distilling out the vitamin A esters, examining the curve with his magnifying glass, and trimphantly announcing that the victim was a Chinese coolie.

The distillation of vitamin A stored in the liver is an extremely interesting problem, the investigation of which is by no means complete. A few years ago Moore in Cambridge, England [19], showed that vitamin E ingested simultaneously with vitamin A increased the storage of vitamin A in the liver. Harris and Cawley in our laboratory have distilled vitamin A recovered from the livers of two groups of rats, one of which was given vitamins A and E while the other received vitamin A alone. It is significant that when vitamin A is given alone a C_{18} type of vitamin A ester is recovered from the liver. When vitamin E is administered with the vitamin A, the C_{18} esters are accompanied by larger quantities of C_{22} esters of vitamin A, and the total storage in the liver is increased. Since the C_{22} esters evidently contain highly unsaturated acids, it would appear that the vitamin E has acted as a preservative not only for the vitamin A but for unsaturated fats which may also be in the class of indispensable food factors. Harris, with J. Jensen, is pursuing the matter further.

At the present time our laboratory is occupied with studies of the sparing action of the E vitamins on one another and on the A vitamins and other labile factors present in food or the animal body [20]. The work is not directly related to molecular distillation but has been derived from the vacuum still by the logical if luxurious process of following each lead to its conclusion. Another of these leads is the preparation and subsequent examination of pure crystalline vitamin preparations [30] from natural concentrates. Pursuing this line, Dr. J. G. Baxter and Dr. P. L. Harris have shown that *natural* optically active α-tocopherol has nearly but not quite double the antisterility activity of the synthetic racemic compound. The excess in potency of the racemic mixture over the expected 50 per cent is evidently due to the intestinal sparing action of the inactive half on the active half.

Limitations of time and space demand a closing of this scarcely impartial account. We have done less than justice to the pioneers in the art of high-vacuum distillation. The work of Waterman [29] on the drying oils, Dam [21] and Almquist [22] on vitamin K, Fawcett [8], Bancroft [23], Burrows [24], and Hecker [25] on the design of stills has not been mentioned. The reader interested in the history and wider application of the method is referred to other reviews [5] [6] [8] and certain excellent bibliographies [26] [27] [28].

ACKNOWLEDGMENTS

Among my associates I wish to thank Dean Cox, L. L. Layton, and E. S. Perry for the photographs of oil films and colored specimens; and Torkel Korling for some of the excellent views of the factory.

For References see p. 316.

X

PRESENT STATUS OF THE VITAMIN B COMPLEX

By C. A. ELVEHJEM

University of Wisconsin, Madison

THE vitamin B complex is composed of several dietary factors classified as a group because like all vitamins they are required by the body in extremely small amounts and because the individual chemical compounds are soluble in water and present in yeast and liver in significant quantities. The significance of these compounds as vitamins was established through the use of a variety of experimental animals. The rat was extensively used in thiamine and pyridoxine studies, the dog in work on nicotinic acid, and the chicken in the early studies on riboflavin and pantothenic acid. Two of the vitamins, nicotinic acid and choline, were known chemically long before their biological significance was recognized. Riboflavin and nicotinic acid were identified as components of important respiratory enzymes before they were shown to have growth-promoting activity in the animal. Pantothenic acid was studied as a growth factor for microörganisms for several years before anyone was able to demonstrate its importance in animal nutrition. Thus it is often more difficult to devise experimental methods for demonstrating the biological activity of a compound than it is to isolate the compound from an extract known to contain an active substance.

Although these compounds are associated in one group their functions in the living cell may differ greatly and their chemical structures may vary from the simple configuration of nicotinic acid to the more complicated molecular structure of thiamine and riboflavin. At present we must recognize at least a dozen separate compounds and six of these, namely, thiamine, riboflavin, nicotinic acid, pantothenic acid, pyridoxine, and choline are now well known to biochemists and nutritionists. Before turning to the newer factors it might be well to review very briefly the important facts about these better-known components.

THIAMINE

Thiamine hydrochloride is a white crystalline substance possessing a nutlike and salty taste and a yeastlike odor. The compound has the structure shown in Figure 89. In foods and in tissues it occurs both in the free form and as thiamine pyrophosphate or cocarboxylase. In the latter form it functions in the living cell as a coenzyme for carbohydrate metabolism. In a thiamine deficiency the metabolism of carbohydrate is incomplete and pyruvic acid accumulates in the tissues, a condition which is used clinically in determining thiamine insufficiency. A variety of symptoms has been described in the human (1) but in all cases the subjects become depressed, irritable, quarrelsome, uncoöperative, and fearful. For many years the rat assay procedure was the only method available for the estimation of thiamine in foods, but more recently the thiochrome and the yeast fermentation methods have been applied successfully. The optimum daily human requirement is 1.5 to 2.3 mg. Although thiamine is widely ditsributed in biological material, only foods, such as pork, oatmeal, and the whole wheat can be considered good sources.

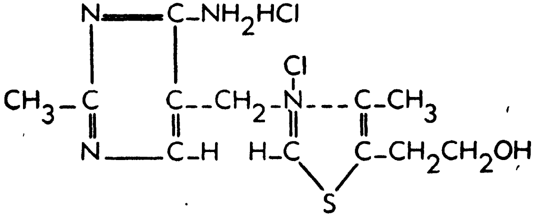


FIG. 89. Thiamine hydrochloride.

RIBOFLAVIN

Riboflavin is a practically odorless, orange-yellow crystalline compound which in water solution shows a characteristic yellow-green fluorescence. It has the structure shown in Figure 90.

In nature riboflavin may occur as such, as riboflavin phosphate, or as a constituent of specific flavoproteins. These flavoproteins function as important enzymes in tissue respiration. Enzymes known to contain riboflavin include cytochrome reductase, d-amino acid oxidase, and xanthine oxidase [2]. A definite reduction in the tissue concentration of the latter enzymes has been observed in riboflavin deficient animals. Aside from retarded growth in young animals, other symptoms include dermatitis and cataracts in rats and characteristic paralysis in chicks. In man the symptoms include inflammation of the lips, fissures of the corners of the mouth, glossitis, dermatitis, and vascularizing keratitis. The riboflavin content of foods may be determined by measuring the growth response obtained in chicks or rats maintained on basal ration low in the vitamin, but the more rapid microbiological method of Snell and Strong [3] is now used extensively. Chemical methods involving measurement of fluorescence are also being used. The daily human requirement is 2.2 to 3.3 mg. Riboflavin is widely distributed in plant and animal materials but liver, milk, and green leafy vegetables are the best and most reliable sources in the human dietary.

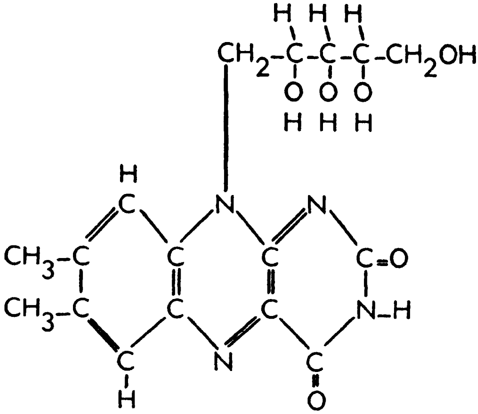
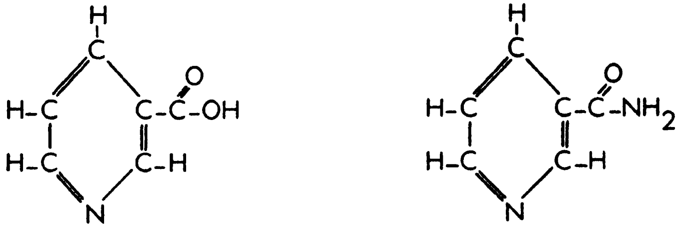


FIG. 90. Riboflavin.

NICOTINIC ACID

Nicotinic acid occurs as white, needle-like crystals, is nonhydroscopic, and is stable in air. The structural formulae of the acid and the physiologically active amide are shown in Figure 91. In the body nicotinic acid functions as a component of two important coenzymes, coenzyme 1 or cozymase and coenzyme 2, both of which are concerned in glycolysis and respiration [4]. In nicotinic acid deficiency it is possible to demonstrate a decreased cozymase content of the liver and muscle tissue.



I. II.

FIG. 91. (I.) Nicotinic acid (Niacin). (II.) Nicotinic acid amide (Niacin amide).

In the early work on nicotinic acid the dog was used extensively for assay purposes. Chemical methods have been used to some extent, but the most satisfactory and convenient method is the microbiological method of Snell and Wright. It has been suggested that 15 to 20 mg. is the daily allowance for an adult man. Liver and yeast are the richest sources of this vitamin although all animal tissues are good sources.

PYRIDOXINE

Pyridoxine was first recognized through its ability to prevent a dermatitis in rats which was observed during attempts to produce experimental pellagra in this species. It was first obtained in crystalline form in 1938 and its synthesis was described by Harris and Folkers [5] shortly thereafter. Pyridoxine hydrochloride is a white, crystalline powder slightly bitter in taste and possessing the structure shown in Figure 92. A deficiency of this vitamin has been produced in chicks, dogs, pigs, and rats. A microcytic hypochromic anemia also results from a chronic deficiency in dogs and pigs. No clear-cut symptoms have been described in man. However, there is no reason to believe that a human being does not require this factor and it has been estimated that the requirement may be 1 to 2 mg. per day. This vitamin has been determined in foods by the rat-growth method and also by a yeast-growth method. This factor is more widely distributed in foods than almost any of the other vitamins.

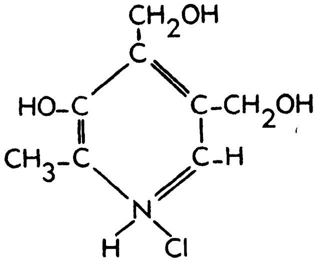


FIG. 92. Pyridoxine hydrochloride.

PANTOTHENIC ACID

The term "filtrate factor" was used for several years to designate that member of the B complex which prevented dermatitis in chicks. Although the so-called filtrate fractions from liver extract were effective in the prevention of black tongue in dogs, pellagra in humans, and dermatitis in chicks, it was recognized as soon as nicotinic acid was accepted as the anti-pellagra factor that the activity of these fractions for the chick was not due to the nicotinic acid present but to a separate and distinct vitamin. Woolley, Waisman, and Elvehjem [6] and Jukes [7] independently demonstrated that pantothenic acid, which Williams [8] had shown to be a growth factor for yeast as early as 1933, was similar to the chick antidermatitis factor. A complete synthesis of calcium pantothenate was achieved by Stiller et al. [9] and the free acid has the structure shown in Figure 93. Rats placed on diets low in pantothenic acid grow very poorly and developed in a few weeks necrosis of the adrenal cortex. When black or piebald rats are used significant changes in hair pigmentation (graying) can be observed. Deficiencies have also been produced in dogs and in pigs. However, little is known about the importance of

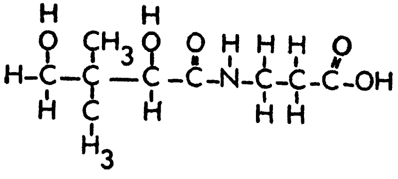
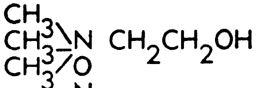


FIG. 93. Pantothenic acid.

this vitamin in human nutrition. If it is needed the daily requirement may fall between 5-10 mg. per day. Meats, cereals and milk are reliable sources of this vitamin.

CHOLINE

Although there may be some question about the inclusion of choline in the B complex, the fact that it is now added to most of the purified diets used in vitamin studies suggests that it is logical to discuss its nutritional significance along with this group of compounds. It has the structure shown in Figure 94. Choline has been recognized for many years as a component part of the phospholipid lecithin, but its possible need in the diet was not apparent until Best [10] demonstrated its role in the prevention of fatty livers in depancreatized dogs. The function of choline is related to the mobilization of fatty acids in the body, since in its absence liver fat accumulates rapidly. The observations of du Vigneaud and his collaborators [11] that the methyl groups of choline as well as those of methionine and betaine are transferable in the animal organism have led to the conclusion that one of the functions of choline is to supply labile methyl groups. Jukes [12] has shown that choline is one of the factors required in addition to adequate manganese to prevent slipped tendons or perosis in young turkeys. Depression of the growth rate when choline is omitted from the diet has been observed in the case of the rat, the chick, and the dog.



$$\begin{array}{l} CH_3 \\ CH_3 \searrow N\ CH_2CH_2OH \\ CH_3 \nearrow O \\ \quad N \end{array}$$

Fig. 94. Choline.

With this background in mind, let us consider how a variety of approaches has helped, and will continue to help, in studies on the remaining members of the B complex. In 1940, when synthetic pantothenic acid first became available, workers in several laboratories were surprised to find that rats grew exceedingly well when this vitamin was added along with thiamine, riboflavin, nicotinic acid, pyridoxine, and choline to synthetic rations free of the B complex vitamins.

Several of us had been working on factors in liver and yeast which appeared to be distinct from pantothenic acid. However, we had been using rather low levels of pantothenic acid, and when the level was increased to 80γ per day, growth of 3.5 grams per day was obtained with the rats. Some workers, such as Richardson (13) and Jukes (14), seemed to think that additional factors were not needed for normal growth, whereas others, notably Woolley (15), Dimick and Lepp (16), Unna (17), and Frost (18), obtained better growth with natural concentrates than with synthetic pantothenic acid alone. There was never any question in my mind about the difference between the rate of growth obtained with pantothenic acid alone and that obtained with pantothenic acid plus natural foods, but the difference was so small that it was difficult to use this growth response as a basis of assay for the remaining factors.

TABLE I

Composition of Purified Ration for Rats

Sucrose	73	gm.
Casein	18	gm.
Salts IV	4	gm.
Corn oil	5	gm.
Thiamine	0.2	mg.
Riboflavin	0.3	mg.
Nicotinic acid	0.25	mg.
Pyridoxine	0.2	mg.
Calcium pantothenate	2.0	mg.
Choline chloride	100.0	mg.

Reproduction in rats receiving only the six B vitamins was first reported by Jukes (14), and reports by Richardson (13) and Unna (17) soon confirmed this finding. We had been obtaining similar results in our laboratory, but it was difficult for me to believe that the rat did not need several additional factors for normal metabolism. In other words, these rations were adequate not because all factors were supplied in the diet but because of the synthetic activity of certain intestinal bacteria. These bacteria were producing the remaining essential factors in sufficient quantities to meet the nutritive requirements of the rat. Work with chicks has shown that these animals cannot be grown on diets similar to those used for the rats; it is generally accepted that the intestinal flora in the rat and chick differ. The guinea pig failed completely to obtain adequate nutrition on diets adequate for the rat, and it has been shown that the number of bacteria in the tract of the guinea pig may be very low.

INTESTINAL SYNTHESIS OF VITAMINS

The possibility of synthesis of vitamins in the intestinal tract is not a new concept. There is sufficient synthesis of vitamin K to make it practically impossible to produce a vitamin K deficiency in the rat, whereas the chick develops a vitamin K deficiency in 8 to 10 days when placed on a diet deficient in this factor. There is also considerable synthesis of vitamin K in the human tract. Examination of the feces of rats maintained on low pantothenic acid diets showed much more of this vitamin than had been ingested with the food [19]. The animals actually excreted 50-75γ per day. However, in this instance the synthesis must have taken place in those areas of the tract where the absorption was very small, because these rats actually died of pantothenic acid deficiency while excreting the vitamin in the feces. The entire intestine from the stomach to the cecum contained 1.5 to 5γ of pantothenic acid, but the cecum and colon contained 25 to 40γ. Synthesis of vitamins in the rumen of the cow is also well established. As early as 1915 Theiler, Green, and Viljoen [20] suggested that "vitamin B" was synthesized

TABLE II

Synthesis of B Vitamins in the Rumen of a Cow

	γ per gram dry matter	
	Synthetic ration	*Rumen contents*
Thiamine	0	10–12
Riboflavin	0.4	18.6
Pyridoxine	1.0	7.0
Pantothenic acid	3.4	55.5
Niacin	1.0	73
Biotin	0.018	0.087

in the alimentary tract of ruminants. Workers at Pennsylvania State College have demonstrated actual synthesis of some of the factors, and at Wisconsin we have found 10 to 30 times as much of each of the B vitamins in the rumen contents as in the feed consumed [21].

Thus there was plenty of indirect evidence for the theory of intestinal synthesis of vitamins, but more direct proof was necessary. I was therefore much interested in the report of Marshall and coworkers [22] that the drug sulfaguanidine was an effective agent for retarding the growth (bacteriostasis) of intestinal bacteria. We had tried a number of different compounds in order to produce bacteriostasis, but we had our first success with a sample of sulfaguanidine (Fig. 95) obtained from Dr. M. L. Crossley, Calco Chemical Division of the American Cyanamid Company. When 0.5 per cent of this drug was added to our synthetic ration, the growth of the rats was reduced from an average of 25 grams per week to less than 10 (Table III). Higher levels had a

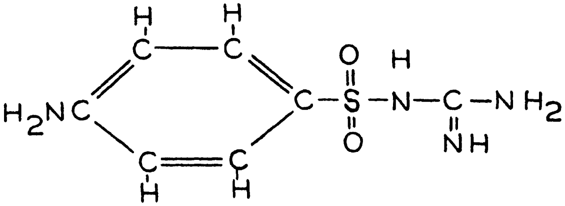


FIG. 95. Sulfaguanidine.

TABLE III

Growth Rate and Prothrombin Time of Rats Receiving Synthetic Ration with 0.5 per cent Sulfaguanidine

	Weekly growth	*Prothrombin time*
	gm.	*sec.*
Control—no sulfaguanidine	25.0	38
Control—with sulfaguanidine	8.7	113
0.3 gm. liver extract	32.0	38
3.0 mg. p-aminobenzoic acid	22.4	47
300γ p-aminobenzoic acid	22.4	43
50γ p-aminobenzoic acid	14.1	105
1 mg. vitamin K per K ration	93.3	40

toxic effect when used with the synthetic ration but gave no inhibition of growth when added to a diet of natural foods. When rats receiving the synthetic sulfaguanidine ration were given 0.3 gram of liver extract per day the growth reached 32 to 35 grams per week. This growth was better than that on the basal ration alone, indicating the importance of additional factors in the liver extract. p-Aminobenzoic acid also counteracted the sulfaguanidine, but the growth was never greater than that obtained on the basal ration without the sulfaguanidine [23]. Mackenzie [24] and others confirmed these results. We assumed that the p-aminobenzoic acid merely counteracted the bacteriostatic effect of the sulfaguanidine and allowed the bacteria to proceed with their synthetic activity, while the liver extract supplied both p-aminobenzoic acid and some of the growth factors used directly by the rat. This idea was substantiated by the fact that rats undernourished on the sulfaguanidine ration showed an immediate growth response to liver but a delayed response to p-aminobenzoic acid.

Since vitamin K was known to be supplied to the rat through the intestinal flora, further verification of our theory would be available if an increased clotting time for the blood from the rats receiving sulfaguanidine could be demonstrated. We were unable to find any difference in the total clotting time, but when the prothrombin time was measured significant differences were observed. It is well known that the prothrombin level can be considerably reduced before a lengthening of the clotting time of whole blood is detected. The prothrombin time of the blood from rats receiving sulfaguanidine was increased from a normal of 40 seconds to over 100. The administration of vitamin K brought the clotting time back to normal, but this treatment had no effect on the growth of the animals. p-Aminobenzoic acid and liver extract counteracted the effect of sulfaguanidine on growth and on the clotting time as well. Further work by Black in our laboratory showed [25] that the effectiveness of the liver extract could also be related to a fraction distinct from p-aminobenzoic acid. The properties of this additional factor were found to be very similar to those reported for an essential growth factor of *Lactobacillus casei* [26] and which has been named folic acid by Williams and coworkers (Table IV).

TABLE IV

Effect of Liver Fractions on Growth of Rats Receiving Synthetic Ration Plus 0.5 per cent Sulfaguanidine

	Weekly growth
	gm.
0.3 gm. liver extract	32.1
Neutral autoclaved L. E	32.9
Acid autoclaved L. E.	14.3
Norite filtrate L. E.	17.1
Norite eluate L. E.	28.2
Acid ether extract L. E.	14.4
Acid ether residue L. E.	22.1
0.6 gm. grass juice	29.6

EFFECTS OF SUCCINYL SULFATHIAZOLE

While this work was in progress Dr. A. D. Welch, of Sharp and Dohme, Inc., made a new drug, succinyl sulfathiazole, available for our investigations (Fig. 96). It also reduced the growth rate of rats on the synthetic diet and caused changes in the prothrombin level of the blood similar to those observed with sulfaguanidine, and these effects were counteracted in a similar way by vitamin K and liver extract. Welch reported [27] identical results last year and added the fact that p-aminobenzoic acid did not counteract succinyl sulfathiazole to the same extent that it did sulfaguanidine. We have verified this observation and have continued to

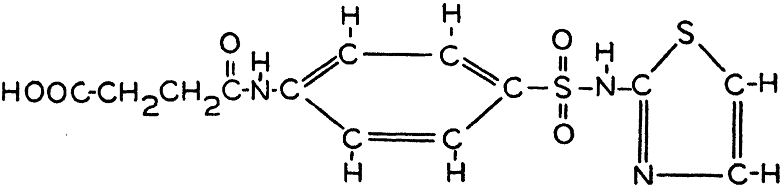


FIG. 96. Succinyl sulfathiazole.

use succinyl sulfathiazole, because the counteracting effect of liver and other materials, such as yeast and grass, can be measured without interference from the p-aminobenzoic acid which may be present. Nielsen and I continued these studies and showed [28] that the reduced growth could be prevented by the addition of biotin and a concentrate of the *L. casei* factor. Finally it was shown that a concentrate of fairly pure folic acid (supplied by Mitchell, Snell, and Williams), given together with biotin, produced good growth in rats receiving the synthetic ration. We now have results to show that liver extract plus biotin gives better growth than a folic acid concentrate plus biotin, even when the concentrate is fed at relatively high levels.

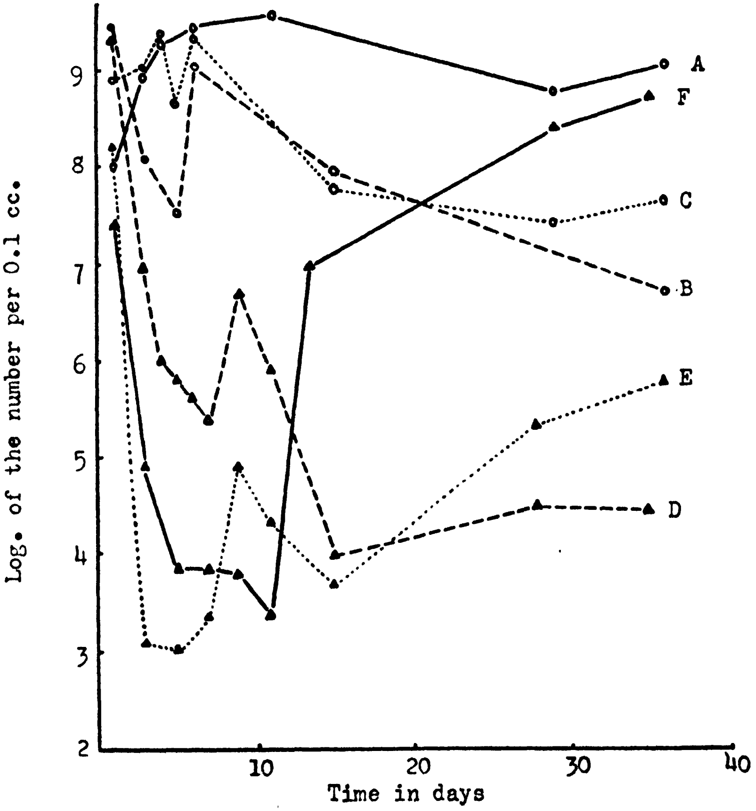
So far I have given only indirect evidence for the changes in the intestinal flora. However, Miss Gant has carefully measured the changes in the bacteria, present in the various segments of the intestines, from the rats receiving the synthetic diet with and without the drug. There is no significant change in the stomach and upper small intestines, but in the lower part of the tract the coliform organisms disappear almost completely upon feeding the succinyl sulfathiazole. The decrease is observed in about 5 days, and the low level is maintained for about 10 days, after which there is a gradual increase in the number of *E. coli*[1] (Fig. 97). This increase is not accompanied by an increased rate of growth; thus the organisms that have adapted themselves to the new intestinal conditions do not synthesize the factors needed by the rat. The total count does not change, which means that other forms can compensate for the reduction in the number of *E. coli.* Poth and others have reported [29] a striking reduction in the coliform count in the feces of dogs receiving succinyl sulfathiazole. The rats receiving the drug plus liver extract or a folic acid concentrate also showed the drop in the *E. coli* count in the colon, and the count remained low for longer periods than without the liver. In spite of this reduced count the rats grew rather well. Thus the folic acid or some other factor in the concentrate is supplied directly to the rat, which can then grow without the aid of these bacteria.

---

[1] *Escherichia coli,* the colon bacillus.

BIOTIN

Biotin was used along with folic acid in these studies because some of the rats receiving the drug showed spectacled eyes, a condition that we had observed frequently in other studies. This condition first interested us in 1939 when rats on low pantothenic acid diets showed this typical dermatitis around the eyes. After the higher levels of pantothenic acid were used this syndrome was never seen. How-



Fɪɢ. 97. Bacteria (*E. coli*) in cecal contents as follows: A, basal only; B, basal plus liver extract; C, basal plus folic acid concentrate; D, basal plus succinyl sulfathiazole and liver; E. basal plus succinyl sulfathiazole and folic acid concentrate; F, basal plus succinyl sulfathiazole.

ever, it was encountered again when we were varying the ingredients of the rats' diet in an attempt to produce graying of the hair in the presence of the higher levels of pantothenic acid. When 10 per cent of the casein was replaced by an equal amount of egg white, the spectacled eye condition appeared uniformly after 5 to 6 weeks.

This result immediately suggested that biotin might be involved, since du Vigneaud, Melville, György, and Rose had shown [30] it to be the active agent which counteracts the injurious effect of raw egg white. In the earlier reports on egg white toxicity we found spectacled eye included as one of the symptoms. When we added a biotin concentrate equivalent to 1γ of biotin per rat per day the eye condition showed great improvement within a week's time. If the animals were continued on the egg white rations for a longer period of time a spastic paralysis developed which would also respond to biotin therapy. When attempts were made to produce similar symptoms by feeding young rats on diets from which the last traces of biotin were removed all the animals grew normally. Thus the rats need biotin, but the required level is produced by bacterial synthesis. Determination of the biotin content of the feces from rats on different rations gave further verification for this idea [31]. The average daily synthesis was about .5γ Sulfaguanidine reduced the rate of synthesis. Diets low in riboflavin and pantothenic acid also reduced the amount of synthesis. This probably explains the appearance of the spectacled eye condition in our earlier studies when diets low in pantothenic acid were used. Daft, Ashburn, and Sebrell have also reported [32] a biotin deficiency in rats that have been given succinyl sulfathiazole. It is interesting that we have been able to demonstrate an increased growth in rats receiving this drug when biotin is given in addition to liver extract. Martin has also reported [33] that biotin and folic acid are growth factors for rats on synthetic diets containing sulfaguanidine, and has concluded that folic acid is a hair pigmentation, or chromotrichial, factor for these rats.

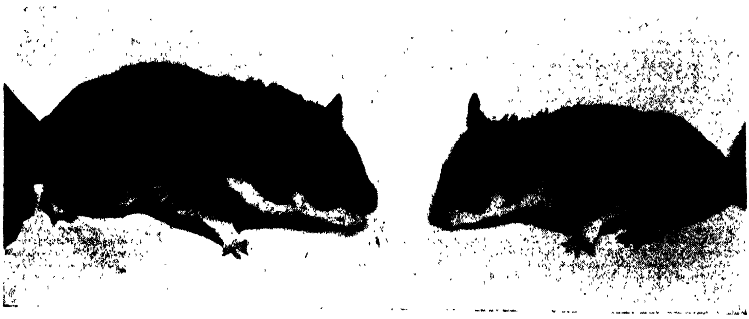


FIG. 98. Effect of biotin deficiency on the growth of the Hamster. Animal on right received synthetic diet adequate in all known vitamins except biotin. Animal on left received same diet plus biotin.

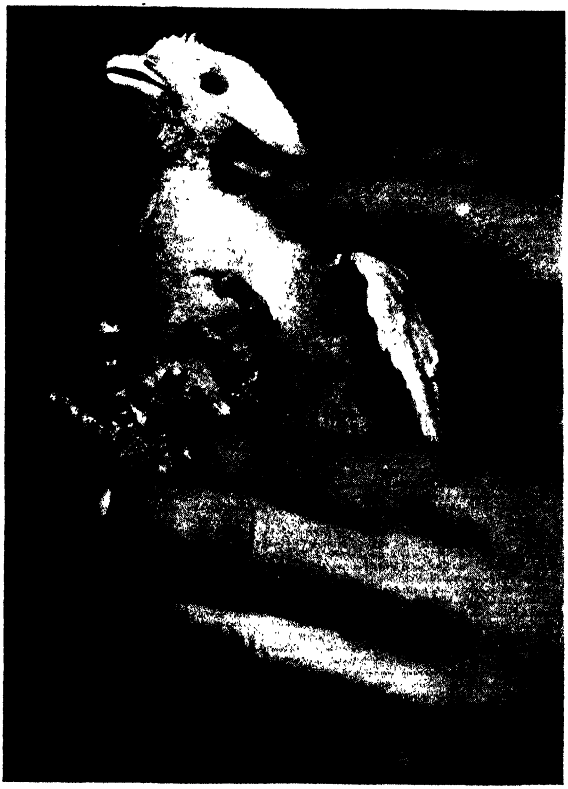


FIG. 99. Dermatitis in chick produced by biotin deficiency.

Two years ago we showed [34] that chicks placed on diets adequate in all respects except for the absence of biotin develop typical dermatitis (Fig. 99). The condition was prevented and cured with biotin. This was the first demonstration of the true nutritional significance of biotin and showed that it is not synthesized in the chick. It is also noteworthy that the Syrian Hamster fails to grow on synthetic diets until biotin is added (Fig. 98). Perhaps the proper flora will develop in the hamster following its more intensive use in the laboratory. From all these results it is clear that biotin is a definite nutritional factor for animals. The formula for biotin as established by du Vigneaud and coworkers is given in Figure 100.

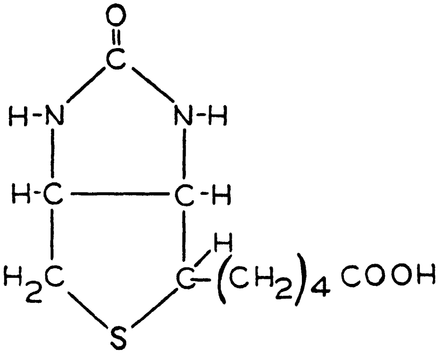


FIG. 100. Biotin.

At this point I want to introduce another bit of evidence for intestinal synthesis — in this instance the amount of synthesis depends upon the level of thiamine in the diet. Much time has been spent on the development of a ration which would be found suitable for the bioassay of thiamine. Various sources of the unknown factors have been used, including autoclaved yeast and peanuts and sulfited liver extract. Due to the good growth obtained when the six synthetic vitamins were used alone, it was suggested that no extra supplements were needed to complete the diet. However, the good growth was always obtained when the six vitamins were given in excess. Our investigations have shown that the level of thiamine in the diet has a profound effect on the requirement for the extra factors.

INOSITOL

The significance of inositol in animal nutrition was first recognized through studies on the mouse. Eastcott showed, as early as 1928, that inositol would stimulate the growth of yeast. In 1940 Norris and Hauschildt found [35] that mice failed to grow on a synthetic diet similar to those we have been discussing, and, in addition, loss of hair and a scaly dandruff were in evidence. The addition of liver and yeast to the synthetic diet produced normal animals. Woolley described [36] a similar condition and identified the anti-alopecia factor as inositol. Further studies [37] indicated that some of the mice, which without inositol showed spontaneous recovery, yielded organisms that would synthesize much more inositol than would the organisms in cultures taken from the tract of mice that remained hairless. This intestinal synthesis was not observed when pantothenic acid was absent from the diet. In the rat there must be sufficient synthesis to meet this requirment. We have recently found that one species, the cotton rat, needs inositol preformed in the diet.

The chick is probably the best animal for studying the additional factors in the B complex. Many difficulties have been encountered in attempts to grow chicks on synthetic diets, but some of these difficulties have now been solved.

TABLE V

A Purified Ration (486) which is Adequate for Chicks when the Liver Fraction which Supplies at Least Two factors, is added

Dextrin	61 gm.	Riboflavin	0.6 mg.
Casein (extracted)	18 gm.	Pridoxine	0.4 mg.
Gelatin	10 gm.	Pantothenic acid	1.5 mg.
Salts V	6 gm.	Choline	150 mg.
Soybean oil	5 gm.	Nicotinic acid	10 mg.
Superfiltrol eluate of solubilized liver	5 gm.	Biotin (concentrate)	0.015 mg.
		Vitamin K	0.5 mg.
Cystine	300 mg.	α-Tocopherol	0.3 mg.
Inositol	100 mg.	Vitamin A	3 drops of conc. per week
Thiamine	0.3 mg.	Vitamin D	

It is now definitely established that the protein requirement of chicks is more exacting than that of mammals. On a diet containing 18 per cent of casein as the source of protein, 3 additional amino acids, namely, glycine, cystine, and arginine, must be added. They can be supplied in experimental diets by using 10 per cent of gelatin and 0.3 per cent cystine in the ration (Table V). Even when these supplements are supplied the chicks fail to grow on the synthetic diet with biotin and inositol added to the first six B vitamins. They also show very poor feathering and a fairly intensive anemia. Fortunately all these deficiencies can be counteracted by adding 2 per cent of liver extract or 5 per cent yeast. The active factors can be collected on the adsorbent, norite, from an acid solution of either of these materials and washed out with a mixture of water, alcohol, and NH_3. Stokstad and Manning first described (38) an active chick factor in yeast which could be adsorbed on fuller's earth and eluted with alkali; they named it factor U. Folic acid and the *L. casei* factor are also present in this fraction, and later we showed (39) that further purification, according to the method used for concentration of folic acid, gave a fraction that was completely active for the chick. However, additional work has shown that the chick requires two factors: one for growth and one for feathered production, both of which are apparently separate and distinct from folic acid (Table VI).

TABLE VI

Growth and Feathering of Chicks Compared to Folic Acid Content of Supplements

Group	*Supplement to basal ration 484 K*	*Weight at 4 weeks*	*Folic acid supplied by supplement*	*Feather development 0 = very poor 100 = very good*
		gms.	γ/100 gms.	
1	No supplement	151	0	33
2	2% solubilized liver	277	50.0	100
3	1% whole liver substance	211	55.3	53
4	Prep. 30 5% sol. liver	240	6.3	85
5	Prep. 33 5% sol. liver	200	16.3	30
6	Prep. 67 5% sol. liver	143	13.8	35
7	Prep. 72 10% sol. liver	185	25.0	35
8	Prep. 73 10% sol. liver	183	5.3	35

Whether folic acid is actually required cannot be answered until the other two factors are obtained free from folic acid.

It should be mentioned here that these factors have now been concentrated into a fraction from liver extract, which is very low in nicotinic acid. Using this preparation in a synthetic ration, containing all the known factors except nicotinic acid, it has been possible to demonstrate that the chick needs nicotinic acid just as do man and dog (Fig. 101). Interesting changes in the nicotinic acid content of various tissues have been obtained by feeding different levels of nicotinic acid.

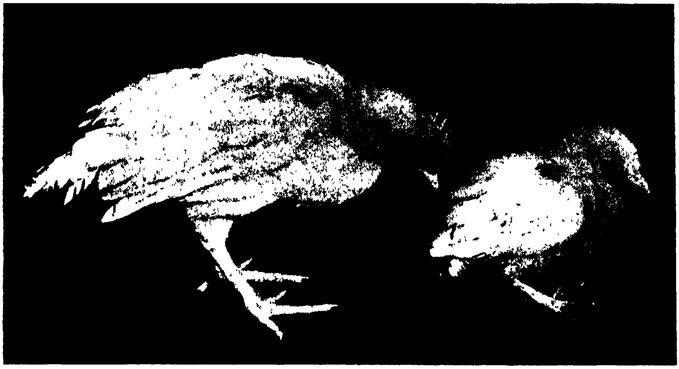


Fig. 101. Effect of nicotinic acid deficiency on chick development.

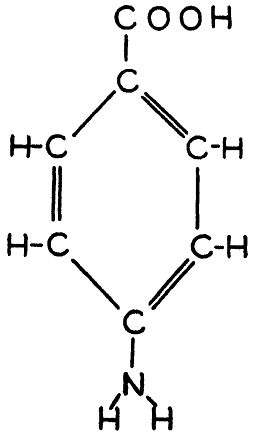


Fig. 102. Para-aminobenzoic acid.

PARA-AMINOBENZOIC ACID

So far I have said nothing about para-aminobenzoic acid (Fig. 102) as a specific growth factor. It was first described as a bacterial growth factor by Rubbo and Gillespie, 1940 [40], and Ansbacher described it [41] as a vitamin in 1941. Sieve has described it [42] as an achromotrichia factor. We have been unable to demonstrate any definite effect of this compound in the rat, except its counteracting effect on sulfaguanidine. Extensive studies made on the chick show that this species can partially compensate for the lack of the liver extract factors when fed at very high levels

(Table VII). However, the liver extract alone, which supplies only very small amounts of p-aminobenzoic acid, gives normal growth and development. Thus if p-aminobenzoic acid is needed in addition to the other factors it is required in only very small amounts. Martin has reported [33] somewhat similar results in the rat. Also the hamster will not grow on the synthetic diet with added biotin unless p-aminobenzoic acid and inositol are supplied but so far we have not determined whether only one or both of these factors are needed.

TABLE VII

Para-aminobenzoic Acid in Chick Nutrition

Supplement to basal ration 486 K per 100 gm.	*Number of chicks*	*Average weight of 4 weeks*	*Feathering 0 = very poor 100 = very good*
No supplement	36	158	32
5 mg. p-aminobenzoic acid	6	146	40
10 mg. p-aminobenzoic acid	24	202	47
15 mg. p-aminobenzoic acid	6	199	65
20 mg. p-aminobenzoic acid	6	167	30
2% solubilized liver	24	238	89

Finally, I should like to mention briefly the guinea pig factors. We have found [43] that this animal will not grow on a synthetic diet until factors present in grass, milk, and yeast are added. Recently Woolley has shown [44] that linseed oil meal carries important nutritive factors for the guinea pig. However, even after that has been added our results show that grass or liver extract and milk are needed for optimum performance. Hogan has reported [45] that guinea pigs do rather well when 15 per cent of yeast is added to a synthetic diet containing dextrin. A few studies made in our laboratory indicate that this response may also be related to bacterial synthesis, since the use of sucrose, in place of dextrin, or the addition of sulfaguanidine retards the growth rate to a very great extent. Mannering has shown [46] in studies on the rat that dextrin allows a greater synthesis of riboflavin in the alimentary tract than does sucrose.

IMPORTANCE OF VARIOUS FACTORS IN HUMAN NUTRITION

The experiments which have been reviewed briefly above involving a variety of animals have demontsrated vitamin-like properties for (1) biotin, which is now available in pure form; (2) inositol and p-aminobenzoic acid, two known chemical compounds; (3) folic acid, which is being studied extensively as a growth factor for bacteria; (4) two chemically unknown factors needed by the chick; and (5) one or more factors of significance in guinea pig nutrition. What is the importance of these factors in human nutrition?

Rather definite information is available for biotin. Sydenstricker and coworkers produced [47] a vitamin deficiency in man by feeding egg white, and then cured the condition with a biotin concentrate. About 30 per cent of the total calories were supplied by the desiccated egg white. Symptoms of dermatitis developed as early as the third to fourth weeks, and many symptoms were similar to those observed in thiamine deficiency. This is an interesting result in view of the fact that we have found a disturbed pyruvic acid metabolism in the livers of biotin-deficient rats. All symptoms were cured by the parenteral administration of 150 to 300γ of biotin per day. Oppel has shown [48] that the biotin content of the urine is influenced by the amount of biotin in the diet. Most of the normal subjects studied excreted from 20 to 50γ per 24 hours, and he was unable to find a single patient who did not excrete biotin Syndenstricker's patients receiving egg white showed a level as low as 3.5γ per day. Oppel also reported that diets of average normal composition contained 30 to 40γ of biotin per day or from 10 to 16γ per 100 grams of dry food. The latter value is interesting, because we found [49] that the biotin requirement for chicks was 7 to 10γ per 100 grams of ration. The total biotin output in humans was found to be 3 to 6 times as great as the intake from the diet. From this it is apparent that an intestinal synthesis of biotin occurs in man as well as in the rat. Since the tract contains large amounts of *E. coli,* it is noteworthy that Landy and Dicken [50] have shown that this organism, along with several others, can synthesize relatively large quantities of biotin. Thus the intake of biotin in man may not be highly important except in special cases. However the lack of some of the other B vitamins may cause nutritive deficiencies due to a reduced synthesis of biotin as well as the deficiency due to a direct lack of the vitamin in the body tissues.

Little can be said about inositol in human nutrition. The fact, however, that various animals require this compound makes it necessary to consider its possible essential nature in humans. Folch and Woolley have isolated [51] a phospholipid containing inositol from brain tissue. This compound is probably present in human tissues and, if it is, the inositol must be supplied either through diet or through intestinal synthesis. If inositol is synthesized there may be some question about its availability. In this connection the recent work of Mitchell and Isbell [52] should be mentioned. Working on cecal bacteria from rats and on the vitamin content of the medium surrounding the bacteria, they found that inositol, nicotinic acid riboflavin, and thiamine diffuse into the surrounding medium to a relatively small extent, pantothenic acid and folic acid to a larger extent, while biotin and pyridoxine move freely from the cells into the medium.

p-Aminobenzoic acid can undoubtedly produce certain effects in man, but I am inclined to believe on the basis of our work with chicks that the action is indirect. Martin has recently taken a similar point of view [33].

Folic acid has taken on new significance in relation to human nutrition, due to indirect evidence from a number of different experiments. First Waisman and Rasmussen, working in our laboratory, have shown that monkeys fail completely to maintain adequate nutrition on synthetic diets similar to those described, even if the diets contain inositol, p-aminobenzoic acid, and biotin. On the other hand, the monkeys live and develop normally if liver extract, grass juice powder, or a crude folic acid preparation is used. It is possible to conclude from this result that the essential active substance is folic acid, because the work with chicks has shown that this preparation contains several factors. However, in the nutrition of the monkeys a grass juice preparation is as active as liver extract; this is not true in the case of the chick. The monkeys on this synthetic diet also show a leucopenia similar to that described by Day and coworkers [53] in monkeys fed a modified Goldberger diet. These investigators found that yeast and liver contained a factor which produced a normal blood picture; this they called vitamin M. The monkeys on the synthetic rations also show severe dysentery, and the stools showed positive cultures of *Shigella paradysenteriae*. Both the leucopenia and the dysentery respond very well to the feeding of crude folic acid preparations. Saslaw, Schwab, Woolpert, and Wilson have reported [54] somewhat similar results with monkeys on synthetic diets.

Now let us consider again the experimental data from rats receiving sulfonamides. Spicer, Daft, Sebrell, and Ashburn have recently reported [55] the consistent development of leucopenia and agranulocytosis in rats receiving sulfaguanidine and succinyl sulfathiazole in synthetic rations. The number of total leucocytes dropped in severe cases from a normal of 10,000 to less than 1,000. I may add that we have verified these results. Lesions in the blood vessels, voluntary muscles, bone marrow, heart and liver in rats fed 1 per cent sulfaguanidine have also been reported [56]. Some of these effects may be related to the toxicity of the drug, but it is curious that leucopenia develops in rats only by feeding sulfonamides, whereas in monkeys it is due to nutritional deficiency. This relationship is especially interesting in that the condition in both animals can be prevented with a crude folic acid concentrate. The fact that the condition develops in monkeys due to a straight nutritional deficiency at least indicates the importance of one or more of these factors in human nutrition. The possibilities appear highly important. Long has emphasized [57] the importance of the time factor in the small number of human patients that show granulocytopenia after administration of sulfonamides. Is it possible that these people suffered from a partial nutritional deficiency and that after about 20 days the drug precipitated this deficiency into a true granulocytopenia?

We can merely speculate as to the importance of the other factors in human nutrition, but I am willing to wager that equally important relationships will develop. As research in this field continues we may discover nutritional means of handling diseases, which today are uncontrollable, as readily as scurvy, rickets, and pellagra are now controlled.

For References see p. 317.

XI

BLOOD AND BLOOD DERIVATIVES

By EDWIN J. COHN

Department of Physical Chemistry, Harvard Medical School

Our understanding of the circulation of the blood followed the rapid development during the sixteenth and seventeenth centuries of our knowledge of anatomy. William Harvey (1578-1657) studied at Padua where the Chair of Anatomy had been held successively by Vesalius, by Realdo Columbus, by Gabriele Fallopio, and by Fabricius of Aquapendente, the latter carrying on the tradition in Harvey's time. Returning to England, Harvey continued his studies and published his great work *De Motu Cordis* in 1628.

Our understanding of the cellular composition of the blood followed the development of the microscope. Leeuwenhoek (1632-1723), fascinated by the phenomena revealed by the lenses he ground, achieved magnifications with the aid of which he described the blood corpuscles first in the frog and then in man.

Comparably, our understanding of the molecular composition and the multiple functions of the blood is following the beginnings in the nineteenth century and the rapid development in the twentieth of protein chemistry. Just as the development of the microscope made it possible to distinguish the cellular elements of the blood, so advances in the methods of protein chemistry are making it possible to distinguish a large number of substances dissolved in the blood, each with different physico-chemical and physiological properties. The separation and concentration of these substances make possible a new mastery and give promise of a new therapeutic control of the composition and the functions of this complex body fluid.

Blood has always been accorded a unique status in the minds of men. It was early associated with the seat of the soul. The Bible distinguished repeatedly between the tissues and the blood. Association with the spirit of man is further attested by the large number of expressions in which special attributes of men are associated with the nature and quality of their blood. One said of a nobleman or of a criminal that it is "in his blood"; of the calm and calculating that he is "cold blooded"; of the wellborn "blue blooded"; of the closeness of family ties "blood is thicker than water." The scientific evidence, as we shall see when we consider the viscosity of blood, is adequate to prove the latter contention, but despite the great advances that have recently been made, it is not yet possible to distinguish chemical attributes of the blood which could form the basis for the other expressions, residues of the folklore of blood.

When Mephistopheles demands of Faust that he sign the compact pledging his soul in blood, Goethe has Mephistopheles explain to the protesting Faust that "blood is a juice of rarest quality." Goethe's *Faust* appeared at the beginning of the nineteenth century. Marlowe's *Doctor Faustus* goes less far in freeing blood of its biblical and medieval mysteries by ascribing to it peculiar properties as a juice or body fluid.

The mysteries that have always been ascribed to the blood have not been entirely solved by modern science. Blood performs more functions than we have learned to ascribe to the substances in it that we recognize. We now recognize many substances in the blood whose functions are not yet understood. We may hope, however, to resolve these mysteries by increasing our knowledge of this "juice of rarest quality" which is the circulating fluid of the body.

THE RED CELLS AND THE RESPIRATORY FUNCTION OF THE BLOOD

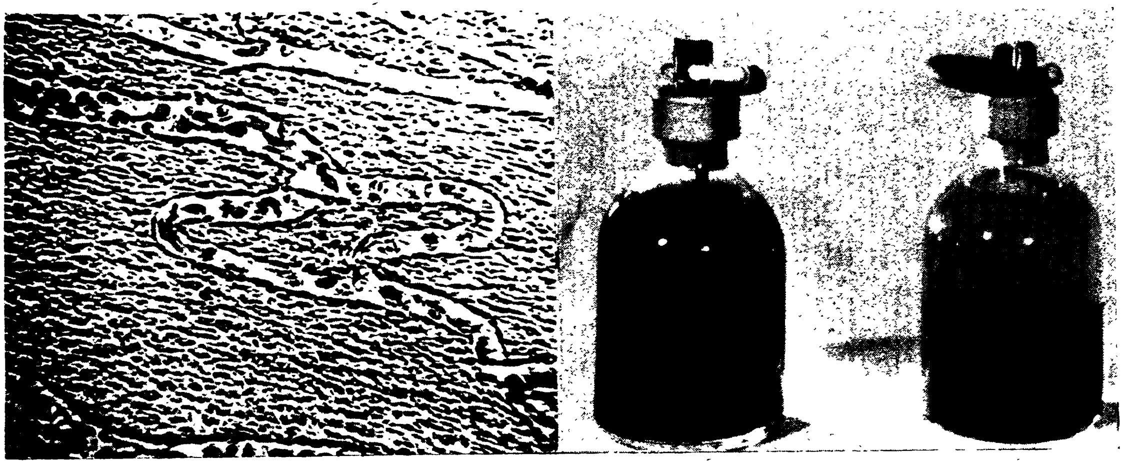
This circulating fluid contains cellular elements of various kinds. Most important among them are the erythrocytes or red cells and the leukocytes or white cells. The red cells perform the respiratory function of the blood; the composition and functions of the white cells are less well known but are generally associated with phagocytic activity.

The circulating blood is in equilibrium with all of the tissues of the body. Anatomists had explored the distribution of arteries, capillaries and veins, and Harvey demonstrated the circulation of the blood propelled by the heart through this complex system. Nineteenth-century physiology recognized and investigated the interchange between the various tissues and the blood which Claude Bernard called the *milieu intérieur,* or internal environment. Considered from the point of view of respiration, the blood carries oxygen to the tissues and removes carbon dioxide and other waste products from them. The mechanism for oxygen transport is not only situated within the red cells but can be associated with certain specific proteins that they contain.

Proteins are among the largest known molecules. All living tissue contains proteins and the complexity of the protein molecule is such that its synthesis in the laboratory is not yet envisioned. However, it is synthesized in processes of biological growth, and the shape and functions of the tissues may be thought of as depending in large part upon the shape and structure of the proteins of which they are composed.

The protein within the red cells which gives blood its color and performs its prime respiratory function is hemoglobin. Proteins are largely composed of nitrogen, carbon, hydrogen, and oxygen organized in the more than twenty different amino acids which are the building blocks of proteins. Hemoglobin contains hundreds of amino acids held together by chemical combination, with loss of water, of the amino and carboxyl groups which all possess. The synthesis of protein in the body is balanced by hydrolysis of these same links which thus confer lability as well as specificity of structure upon these vast protein molecules.

Hemoglobin performs its specific respiratory function by virtue of a relatively small prosthetic group, attached to the molecule but readily separated from it, which is not composed of amino acids but of iron in a configuration of pyrrole rings. The iron in this configuration has a high avidity for



A

B

Fig. 103. A, Blood in capillaries showing red cells magnified 175 times (photomicrograph by courtesy of Dr. E. M. Landis). B, Blood in Red Cross collecting bottles. In the bottle on the right, after centrifugation, the red cells are concentrated at the bottom, the plasma above.

oxygen, combines with it, and is the means of its transport. The prosthetic group containing the iron is called heme, the protein part of the molecule globin.

Let us examine the relative masses of the oxygen needed by the tissues and the various component parts of the mechanism which the body employs for its transport. The atomic weight of oxygen is 16 and of iron 55. Each iron atom combines with an oxygen molecule, that is, two oxygen atoms. The iron, however, is part of the heme group with a combining weight close to 700. The heme is attached to the globin, each iron to a unit of 16,700 equivalent weight. Iron-containing proteins concerned with respiration in muscle, known as myoglobin, are of this size, but the hemoglobin of the blood stream has a molecular weight of at least 33,400 and generally of 66,800, in the former state combining and transporting two, in the latter four, oxygen molecules. The molecule devised by nature for the efficient transport of oxygen to the tissues thus has a mass roughly 500 times that of the oxygen molecule to be transported.

Why is the hemoglobin of the blood stream contained within cell walls? The proteins which perform comparable respiratory functions in many forms of life are free in the fluid part of the blood. The proteins dissolved in the fluid part of the blood of animals rarely exceed 10 per cent, those dissolved in the fluid or plasma part of the blood of man, 7 per cent. This figure may be contrasted with the 30 per cent of hemoglobin contained in the red blood cell. Were this concentration of hemoglobin free in the blood stream, it would bring about a redistribution of many other components than oxygen. It is the way of the body, or appears to be to those concerned with the study of its mechanism, to evolve processes which in enhancing the efficiency with which certain functions are performed do not inhibit others.

The function of transporting oxygen by the hemoglobin within the red cell involves a complex mechanism permitting maximum combination with oxygen in the lungs and release of oxygen in the tissues. As carbon dioxide increases, the hemoglobin unloads the oxygen which it has transported.

Another protein, an enzyme, carbonic anhydrase, accelerates and renders more efficient this process. In addition, the red cells contain other enzymes, including a catalase and a phosphatase. The presence of these enzymes suggests the nature of certain of the chemical reactions which are involved in the internal economy of these cells. These vital processes also appear to limit the life of the red blood cell. For the life of the red cell, estimated to be roughly three months in the blood stream, is limited in the blood bank, under even the best methods of preservation at present known, to approximately a month.

No other part of human blood has so significant a function in transporting oxygen to the tissues, in permitting the very great activity and high order of mammalian respiration. Therapeutically, therefore, when there is great blood loss from a wound, when a major operation is to be performed, or when the anemia which frequently occurs in convalescence from wounds must be combated, red blood cells must be supplied in amounts adequate to restore the oxygen carrying capacity of the blood.

On the other hand, the lability of the red cells renders it difficult to develop a reserve of whole blood which can be transported to the theaters of operation at great distances and stored indefinitely against the uneven needs of a military campaign. There are, moreover, many functions of the blood other than those of oxygen transport. These can be supplied by other parts of the blood, the more labile red blood cells being conserved for therapy for conditions in which they alone suffice.

PLASMA AND THE PLASMA PROTEINS

Blood is readily separated into cellular and non-cellular parts by centrifugation. The heavier cellular elements are easily thrown to the bottom when the bleeding bottle is spun in a cup centrifuge. The supernatant fluid is clear, amber colored, and is called the plasma. The separated red cells occupy something less than half the total volume; the plasma, the body fluid in which they are suspended in nature, something more than half the volume. The proportion of the blood occupied by the red cells varies somewhat from individual to individual. If the red cells are too few in number, the condition of anemia obtains; if they are too concentrated, the viscosity of the blood is increased, and an extra burden is imposed on the heart in forcing the blood through the circulatory system.

Shock as observed in military medicine generally results from a rapid decrease in the volume of circulating blood due to the loss of blood and plasma proteins externally and into damaged tissues. It is most frequently treated, or prevented, by injection of plasma proteins. The injected proteins insofar as they cannot readily pass through the kidney increase the body's reservoir of plasma proteins, and insofar as they do not readily transverse the capillary walls, increase the plasma volume by drawing water from the tissues into the blood stream, and by holding it there.

The maintenance of the fluid balance between the blood and the tissues is one of the functions of the plasma proteins. While they are like hemoglobin in their general organic structure, their special properties and their functions differ widely from hemoglobin or globin. Moreover, there are many different kinds of plasma proteins, each serving a different function. The function of controlling the equilibrium between the water and the electrolytes in the blood and in the tissues is largely performed by the smaller proteins of the blood stream, known as the albumins, although all dissolved colloids will, of course, exert some osmotic effect.

Osmotic pressure is the force exerted by dissolved molecules so large that they cannot pass through the pores of the membranes that contain them. The walls of the blood vessels are the retaining walls for the plasma proteins. Oxygen, water, electrolytes, sugars, and other small molecules readily pass these walls. But normally the plasma proteins do not, nor do they pass the kidney and appear in the urine. A variety of conditions influence the permeability of membranes to dissolved molecules. Besides the size and shape of the molecules, these include the charged state of the membranes as well as of the molecules and the nature of their surface groups. The limiting condition for passage through a membrane is however the size of the pores and the size of the dissolved molecule.

The plasma proteins vary in mass and in length. All that have been adequately studied, however, have closely the same diameter, ranging from 33 to 38 Ångstrom units (an Ångstrom unit being 10^{-8} cm.). The albumin molecules are the most symmetrical of the plasma proteins, being approximately four times as long as their diameters. (See Table I.)

TABLE I

Relative Dimensions of NaCl, Glucose, Hemoglobin, Red Blood Cells, and the Various Plasma Proteins

	Approximate dimensions (Ångstroms)		*Viscosity coefficient*	*Molecular weight*
	Length of ellipsoid	*Equatorial diameter of ellipsoid*		
Sodium ion	1.9	1.9	..	23
Chloride ion	3.6	3.6	..	35.5
Glucose	9.5	6.5	4	180
Hemoglobin	118	28	..	39,000
	158	32	5.1	68,000
Red blood cells	24,000	86,000	4	......
Serum albumin	150	38	5.8	69,000
Serum γ-globulin	320	36	8.5	158,000
Fibrinogen	900	33	70	500,000

Among plasma proteins certain of the globulins are twice as long as the albumins and the protein that forms the matrix of the blood clot, fibrinogen, is a rod-shaped molecule six times as long as albumin. Fibrinogen is almost thirty times as long as its diameter. Its solutions are therefore extremely viscous, viscosity being a function not of the size of molecules but rather of their asymmetry. Fibrinogen is present in the plasma, however, to only 0.24 to 0.3 per cent, representing 4 to 5 per cent of the plasma proteins. By contrast, the albumins represent 55 to 60 per cent of the plasma proteins, are by far the most soluble, and their solutions are the least viscous.

Elementary considerations indicate different ways in which the size and shape of the plasma proteins influence the blood. The osmotic pressure is, as a first approximation, a function of the number of dissolved molecules to which the membrane is impermeable. Since the blood vessels are impermeable to the normal plasma proteins they exert sustained osmotic pressure. Since these proteins cannot pass through the membranes their presence in the blood stream holds water in it in an amount roughly proportional to their concentration, more exactly to their effects on the osmotic pressure.

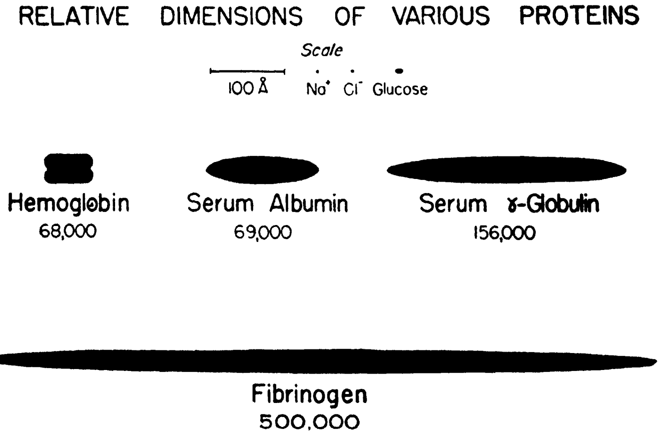


FIG. 104. Relative dimensions of various blood proteins. If the serum albumin molecule were the length of this page, a red cell would be the size of a football stadium. A value of about 6,000,000,000,000 is obtained for the weight of Avagadro's number of red blood cells, which are therefore about a billion times the weight of a serum albumin molecule. Scale models and many of the estimates by Oncley.

Should an increase in permeability of the blood vessels occur which is sufficiently great to allow plasma proteins to escape into the tissue spaces, fluid will also pass into the tissue spaces. Commonly such a change in permeability occurs where there has been a severe burn or injury and usually is not sufficiently great to permit many red cells to leave the blood stream. As a result, there is an increase in red cell concentration (often measured in terms of a hematocrit reading), an increase in viscosity, and a decrease in the volume of circulating plasma. This diminution of blood volume ultimately results in poor peripheral circulation and hence inadequate oxygen transport, while the increased viscosity makes it more difficult for the heart to move the remaining blood through the blood vessels. This condition may be overcome by injecting into the blood stream molecules which will not readily leave it and which, by exerting osmotic pressure, will pull water back into the blood stream, decrease viscosity, and permit a return to normal circulation of the blood with the transport of oxygen to the tissues and the fulfilment of the other physiological functions of the blood.

The diameters of all the plasma proteins are so nearly the same that a change in permeability that will permit one type of protein to pass a membrane will generally permit all to pass. The amounts of each protein type that will pass will, however, now depend upon other factors controlling permeability. Thus the longer molecules will on the average leave the blood stream less readily than the more symmetrical albumins. The electrically charged conditions of the molecules will also play a role. The albumins bear the largest negative charges, and a species of globulins, known as the γ-globulins, precisely because of their low negative electrical charge in neutral solutions, the lowest.

The plasma proteins that leave the blood stream and appear in the urine or in the tissue spaces, under a variety of conditions, might be expected to reflect all of these influences. It has been repeatedly noted by various investigators under various conditions that the distribution of proteins lost from the blood is far more nearly constant than the composition of the plasma under pathological conditions. It is more nearly characteristic of normal than of pathological plasma. Generally it is the albumins and the γ-globulins that appear to be lost more readily than either other globulins of larger net charge or fibrinogen. The significance of this observation is twofold. On the one hand, it reflects the conditions that obtain when proteins leave the blood stream, on the other the replacements that must be made if normal composition of the blood is to be restored by therapy.

Although the smallest plasma proteins of largest net charge, the albumins, and the largest globulins of smallest net charge, the γ-globulins, leave the blood stream most readily, it is the albumins which have the largest effect in controlling the equilibrium in water and electrolytes between blood and the tissues. The γ-globulins represent but 11 per cent of the plasma proteins, the albumins five or more times that much. The molecular weight of the γ-globulins is two or more times that of the albumins. On the basis of mass alone, therefore, the γ-globulins—which as we shall see are the bearers of the immune properties of the blood—contribute less than one tenth as much as the albumins to maintaining the volume of the blood. This calculation neglects the greater contribution of the albumins which depends upon their greater net charge. Indeed, albumins are responsible for nearly 80 per cent of the osmotic activity of the blood. Accordingly it is the albumins in plasma which are in large part responsible for the therapeutic value of plasma in the treatment of shock.

The fact that fibrinogen leaves the blood more slowly than albumin may be associated with its greater length. On the whole, molecules of the same diameter should pass an inert membrane inversely as their length. The presence of fibrinogen in the blood in small amount must be associated with its special functions, which we shall examine later in connection with the clotting of the blood. Its osmotic contribution is negligible but its contribution to viscosity is enormous and this effect might be deleterious were fibrinogen present in large amount. Under these circumstances, it brings about clumping of the red cells as do many other highly asymmetric molecules.

If substances other than the plasma proteins are injected into the blood, their fate also appears to depend upon their dimensions. Saline, glucose, and most substances that have thus far been suggested as blood substitutes have diameters less than 20 Ångstrom units and, unless they are very asymmetrical, leave the blood stream rapidly. They thus exert temporary but do not have sustained influence in maintaining blood volume.

Asymmetric molecules appear to leave the blood stream more slowly and may even accumulate in the blood stream as evidenced by their influence on the clumping and sedimenting of red cells. Such molecules increase the viscosity of solutions and may well bring about abnormal distribution of plasma proteins. It is by no means certain, therefore, that the retention of foreign substances of this kind in the blood stream is beneficial.

From a physico-chemical point of view, it is possible to understand, at least as a first approximation, certain of the phenomena of plasma loss. Since all of the plasma proteins have about the same equatorial diameter, any change in the permeability of a membrane which would permit one plasma protein to diffuse through it should permit any of them to do so. Under these circumstances, the amount of plasma protein of each type to diffuse should be an inverse function of their length. Consequently for a given change in permeability there will be a greater loss—and therefore a greater need for replacement—of albumin than of the globulins or fibrinogen.

Clearly there is a principle guiding the amount of any plasma component that may wisely be used in replacement therapy, namely, the amount that will tend to reëstablish rather than to derange the normal concentrations of the various physiologically significant plasma components.

CHEMICAL FRACTIONATION OF THE COMPONENTS OF PLASMA

The highly specialized structure of the plasma proteins indicates that each has a special function and that their colloidal properties, which alone have thus far been considered, are but superficial manifestations of their intrinsic nature.

The proteins, as their name implies, are "of the first importance," and this importance extends equally to the structure of cells and to the processes of life. The role of these nitrogen-rich substances in the nutritional cycle of plants and animals has long been appreciated. Indeed, the elementary composition of "blood albumin" and "blood fibrin" were being studied by the group of chemists that collected around Liebig just over a century ago, and in 1841 Denis, a French scientist, communicated to Liebig the separation of blood proteins into albumins and globulins.

Though less labile than cells, proteins also are labile. The molecules that occur in nature were until recently not readily separated, purified, concentrated, or dried, without change of structure or loss of function, by the conventional methods of chemistry.

The methods of the nineteenth century have been largely supplanted by recent developments in the preparative chemistry of the proteins and upon two of these are based the preservation and accumulation as against military needs in distant theaters of operation of the plasma proteins separated from the blood collected by the American Red Cross.

The first of these methodical developments is the drying of proteins from the frozen state. Proteins coagulate at elevated temperatures and in the presence of the organic solvents generally used in drying ordinary chemicals. Likewise, denaturing changes often take place when proteins are dried at ordinary temperatures. A body of investigation has demonstrated, however, that such changes are minimized when protein systems are first reduced to the frozen state and water then removed under reduced pressure by sublimation. The dried plasma which has so successfully been used by our Armed Forces in every theater of operation is rendered stable by this technique. The plasma is reconstituted just before use by addition to the dried powder in one bottle of the diluent transported in another bottle in the same package. All but the most labile components of plasma may be preserved by this process, whereas all but its colloidal attributes are lost if liquid plasma is preserved for long periods of time even at low temperatures.

The second development has made possible the quantitatively reproducible fractionation of plasma by new chemical methods which on the one hand are susceptible to the large industrial scale that has been necessary to process the blood

TABLE II

Whole Blood

	Principal function	*Principal protein related to function*	*Concentrated in plasma fraction*	*Plasma protein in fraction*	*Concentration of active function times that of plasma*		*Established clinical use*
				Per cent	*Per gm. protein*	*Per cc. product*	
RED CELLS (45% of blood containing 30% hemoglobin)	Respiratory	Hemoglobin					Whole blood or red cell transfusion.
PLASMA (55% of blood containing 7% proteins)	Osmotic regulation of blood volume	Albumin	V	48	1.3	5.4	Shock, burns, edema, hypoproteinemia.
	Blood coagulation	Fibrinogen	I	6	21	..	Fibrin films as dural substitutes.
		Prothrombin, thrombin	III-2	3	20	..	Fibrin foam and thrombin as hemostatic agent.
	Immunological	Blood grouping globulins	III-1	3	24	36	Blood grouping.
		Complement C'1	III-2	..	15	..	
		Complement C'2	IV	..	8	..	
		γ-globulins	II	6	8	25	Measles prophylaxis.
	Carbohydrate and lipid solution and transport	β globulins	III IV				
		α-globulins	IV				
	Regulatory	Fibrinolytic enzyme	I, III-2	10			
		Phosphatase and other enzymes	III, IV				
		Hypertensinogen	IV-3, 4				
		Thyrotropic hormone	IV-3, 4				
		Gonadotropic hormones	III, VI				

of over one and one half million Red Cross donors, on the other to yield as final products plasma proteins of unimpaired function.

The potential value of all the different proteins of human plasma demanded that this process be inclusive in the sense that all components be separated and concentrated. The present process separates the plasma into five fractions. Fraction I is rich in fibrinogen. Fraction II consists of γ-globulins, which are the chief bearers of immunity in the blood. (For convenience in manufacture, Fraction II + III, which contains essentially all the γ-globulins, is first separated as one and then subjected to subfractionation.) Fraction III-1 contains those euglobulin components of plasma, the isoagglutinins, which interact with the red cells according to their specific blood groups and which, therefore, can be used in the typing of blood. Fraction III-1 is also rich in certain lipo-proteins containing all of the carotenoids and a large part of the cholesterol. Fraction III-2 is largely β-globulin and contains components important in blood coagulation, among them thrombin, prothrombin, and a fibrinolytic enzyme. Fraction IV contains the remaining β-globulins and most of the α-globulins. Of these α-1 is further separated into Fraction IV-1. Fraction V contains between 85 and 90 per cent of the albumin. The mother

TABLE III

*Distribution of Proteins of Plasma**

	Grams protein per liter plasma	*Grams protein per liter of plasma in fractions*				
		I	*II+III*	*IV*	*V*	*VI*
Protein	60.3	4.3	16.3	9.7	29.6	0.6
Albumin	33.2	0.2	0.7	1.0	29.0	0.3
α-Globulin	8.4	0.2	1.8	5.4	0.6	0.3
β-Globulin	7.8	0.8	6.2	3.1	..	..
γ-Globulin	6.6	0.5	6.0	0 2	..	.
Fibrinogen†	4.3	2.6	1.6	..	..	..

* Cohn, Oncley, Strong, Hughes, and Armstrong.

† Estimated by electrophoresis. The amount of clottable protein is always lower and probably more reliable. See Edsall, Ferry, and Armstrong.

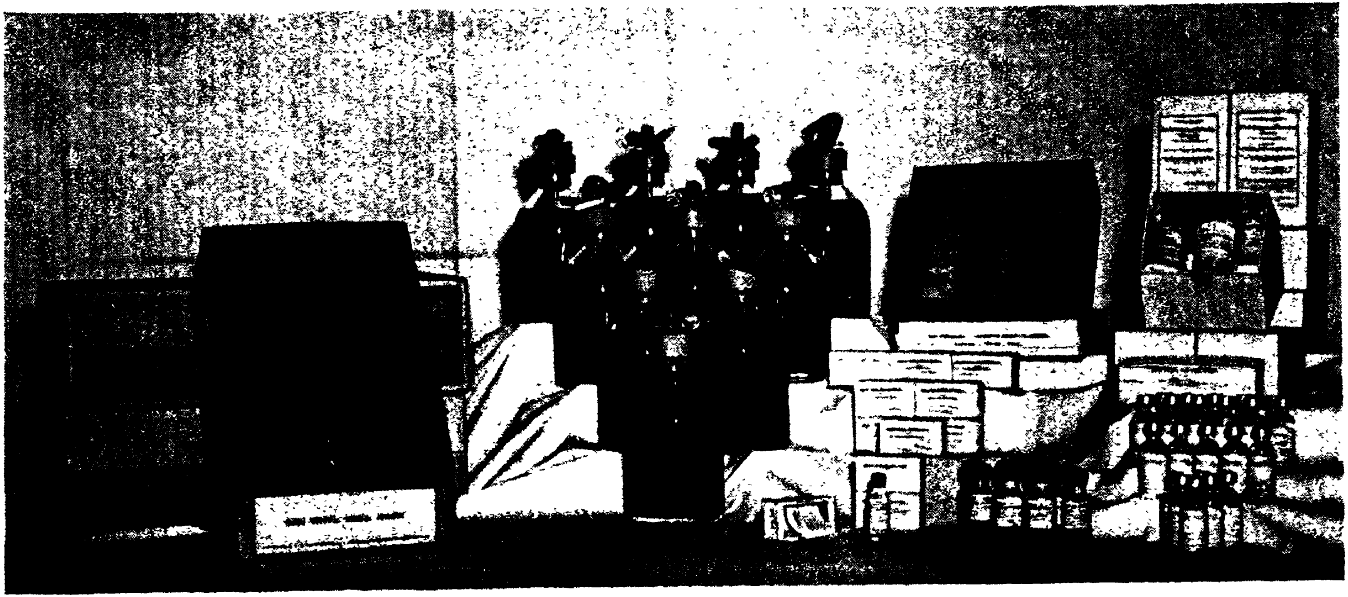


FIG. 105. Products prepared from the Red Cross blood shown in center of picture. The dried plasma obtained from this much blood is shown on the left. The products of plasma fractionation including albumin, serum γ-globulins, fibrinogen and thrombin, fibrin foam and thrombin, and blood typing globulins obtained from the same amount of blood are shown on the right.

liquors in Fraction VI contain about 1 per cent of the total proteins. The yields of the proteins in the various fractions are given in Table III.

The fractionation of plasma into its component parts according to this process developed in the Department of Physical Chemistry of the Harvard Medical School has been carried on in seven commercial laboratories under contract with the United States Navy. The albumin of plasma represents by far its largest component and is concentrated therefore in its largest fraction, V. Whereas there is evidence of the existence of more than one albumin, this fraction appears homogeneous in the ultracentrifuge and in the electrophoretic apparatus at neutral or slightly alkaline reactions. The standard of purity specified in the Navy contracts for the production of this blood derivative permits 2 per cent of globulin. Electrophoretic analyses reveal the constancy of the product under the conditions of industrial production.

Human serum albumin has been still further purified by crystallization, and study of the crystallized albumin that we have prepared has demonstrated that the very high thermal stability and low viscosity of our standard preparations are indeed ascribable to the albumin, instability largely to globulin and lipid impurities.

Normal human serum albumin can be prepared in far larger amounts and with higher yields if further purification by crystallization is not superimposed upon the process. The standards of purity determined upon for the albumin now being delivered to the Navy in large amounts were chosen so as to assure freedom from untoward reactions with maximum efficiency in large-scale production.

Serum albumin was developed in order to attain a blood derivative which could be distributed in solution ready for immediate emergency use. It has been made available as a 25 per cent solution to render the package as compact as possible for transport. Osmotically more than 4 times as concentrated as plasma, 25 per cent albumin is no more viscous than whole blood. The package developed by the Armed

Forces for the standard 25 per cent albumin solution occupies one sixth the space of the osmotically equivalent plasma package.

ALBUMIN AND THE CONTROL OF THE BLOOD VOLUME

That the value of plasma in controlling the volume of the blood depends upon its osmotic activity is beautifully demonstrated by the studies that have been carried out on the osmotic pressures of plasma and of albumin solutions. These measurements yield a molecular weight of 69,000 for albumin and an average weight of about 170,000 for the 40 per cent of the protein which is not albumin. Each gram of albumin is equivalent to 1.2 grams of plasma protein or to 20 cc. of the current Red Cross citrated pooled plasma. The volume of fluid held in the blood stream by each gram of albumin should be about 18 cc.

TABLE IV

Osmotic Effect of Concentrated Human Serum Albumin in Increasing Blood Volume

Conditions of measurement	*Increase in blood volume per gram albumin*
	cc.
Estimated from osmotic pressure*	18
In experimental hemorrhage†	17
In clinical shock, after ½–1 hour‡	16
In clinical shock, after 1–36 hours§	23

* Scatchard, Batchelder, and Brown.
† Heyl, Gibson, and Janeway.
‡ Warren, Stead, Merrill, and Brannon.
§ Cournand, Noble, Breed, Lauson, Baldwin, and Richards.

The results of these osmotic pressure measurements may be compared with the clinical findings of three groups that have made independent studies on the value of albumin for transfusions, especially in shock. All have determined the water drawn into the blood stream following injection of albumin in man. One group reported that "the average increase" in blood volume (one hour after injection) "was 17.4 cc. per gram although there was considerable variation between individual experiments." Another reported that "the albumin injected into the circulation is largely retained, providing no further bleeding or exudation of plasma occurs. It is still remaining in the circulation at the end of 6 hours. The albumin retained holds in the blood stream amounts of fluid approximately comparable to its osmotic activity. Thus, in our series, 1 gram of albumin retained resulted in an average increase of plasma volume of 23 cc." The third group concluded that "the plasma volume increased 600, 800, and 900 cc. in the patients receiving 50 grams of albumin. The average increase in volume was 16 cc. for each gram of albumin." These three carefully controlled clinical investigations yielding respectively 17.4, 23, and 16 cc. drawn into the blood stream of patients per gram of albumin injected must be considered in excellent agreement with the expectation from the osmotic pressure measurements of 18 cc. per gram of albumin which forms the basis for the osmotic equivalence of the two blood derivatives, plasma and albumin, being dispensed to the Armed Forces.

α- AND β-GLOBULINS AND THE HORMONES, ENZYMES AND LIPO-PROTEINS OF THE BLOOD

Although the albumins of plasma play far more important roles than the globulins in the control of blood volume, the diverse globulins also perform functions for which they appear to be uniquely designed. The number of such globulins is far larger than our present physico-chemical criteria for their characterization. Moreover, certain of them appear to be present in very small amounts, amounts, however, adequate to perform their natural functions. In diseases in which there is a deficiency of any of these proteins, however, it would appear to be of the utmost value to have them available, separated from the plasma and concentrated as specific therapeutic agents.

The blood of man is rich in lipid as well as in protein material. Only a small amount of this lipid is present in an emulsified state; most of it is held in solution in more or less labile combination with the various plasma proteins. Among the lipids that are readily distinguished are cholesterol, carotenoids, phosphatides, and cerebrosides. Of these the carotenoids are found concentrated with certain β-globulins in Fraction III-1 (see Table II) which is also rich in cholesterol. Cholesterol, phosphatides, and cerebrosides have also been found concentrated with α-globulins in Fraction IV. Here our study of plasma proteins is leading us to a study of the plasma lipids, not as organic chemicals separated from blood by methods which destroy the structure of the lipo-proteins but in the state in which they exist in nature. Their separation and purification as protein complexes should make possible investigation of their function and an increasing understanding of their role in the economy of the body.

The economy of the body is controlled by a group of messengers, hormones, secreted and liberated by various glands. These hormones also find their way into the blood stream and some of them have been identified, although they are present in but very small amount. Among them is the thyrotropic hormone which appears to be closely related to that secreted by the anterior pituitary gland. The gonadotropic hormones secreted by this gland have also been identified in one or another plasma fraction. As our survey of the hormones in the blood expands it should be possible not only to concentrate and recognize more of them but also to make them available as concentrates should they prove to have any value in therapy. Their human origin should make it possible to use hormones derived from man when those derived from animals lead to sensitization following repeated injection.

The economy of the body is aided also by a series of enzymes whose specific function it is to accelerate chemical reactions. The blood stream contains many such enzymes. A cholinesterase and phosphatase clearly have functions related to the cerebrosides and phosphatides of the blood stream. Phosphatides of more than one kind have been recognized both within and without the red cell. There are lipases and proteolytic enzymes and the number of these is not yet clearly defined. One, the fibrinolytic enzyme, will serve to illustrate a principle that applies to others. The blood stream carries not only components necessary to form a clot instantly at a bleeding surface, but also this enzyme capable of dissolving the clot.

FIBRINOGEN, THROMBIN, AND THE CLOTTING OF THE BLOOD

The clotting mechanism of the body is indeed complex. The structural element of the clot is the long rod-shaped fibrinogen molecule which is dissolved in the plasma. Its aggregation and separation as the fibrin clot are carried out by the globulin, thrombin. When thrombin and fibrinogen are mixed *in vitro*, a clot immediately occurs. Thrombin however does not exist free in the blood stream but as a precursor substance, prothrombin, readily transformed into thrombin in the presence of calcium by tissue extracts known as thromboplastins.

The various parts of this complex mechanism may be separated and concentrated and made available in various forms. By controlling the conditions of the reaction between fibrinogen and thrombin, structures of a wide variety of properties can be obtained. Clots of one type are clear, gelatin-like, elastic, and friable, do not readily synerize with loss of fluid, and adhere to surfaces on which they are formed. Clots of another type are opaque, doughy, and scuff-resistant, synerize readily, become dense with loss of fluid, and do not adhere to underlying surfaces. Clots of intermediate properties can also be made.

For surgical use in human patients, clots may be prepared from fibrinogen and thrombin to meet the particular specifications chosen for each type of application, and may be formed *in situ* in or over wounds. Clots whose properties have been controlled in this way have been used in pyelolithotomy and in a technique for skin grafting.

Just as the asymmetry of certain synthetic polymers is responsible for the remarkable mechanical properties of industrial plastics, the fibrinogen molecule endows products of this human protein with remarkable mechanical properties. Among these products is fibrin film. Fibrin film may be prepared from fibrinogen and thrombin as a strong, rubbery sheet which can be stretched reversibly from two to three times its original length. This material can be made in various shapes and thicknesses, and in the form of seamless tubing. Its mechanical properties can also be varied from a soft, rubberlike elasticity to a parchment-like consistency.

The mechanical properties of the soft, rubbery films (as shown by their stress-strain curves) bear a marked similarity to those of the elastic ligament of the neck, the *ligamentum nuchae*. The tough films with delayed elasticity resemble rather wool or hair in their mechanical properties. It may be possible to imitate, with different types of fibrin products, still other natural structures of the body.

The fine structure of fibrin film has been shown to involve pores which in one type are of the order of 60 Ångstrom units in diameter. Hemoglobin molecules in solution pass through these pores readily, but plasma globulins are partially and fibrinogen molecules completely retained.

The mechanical properties of certain types of fibrin films make them suited for use as dural substitutes and in the prevention of meningocerebral adhesions. The duration of their persistence in the body can be adjusted by suitable treatments. They have been used in neurosurgical operations and appear to be an excellent material for these purposes, patients having been followed for as long as fifteen months without the appearance of unfavorable sequelae. Other types of fibrin film and of fibrinogen plastics are being tested for other surgical applications.

For stopping the flow of blood, in surgical procedures, thrombin is the only component of the clotting mechanism which must be supplied. For the most effective use of thrombin in hemostasis, however, it must be applied with a matrix which can hold the thrombin in the bleeding area until clotting is completed. A porous matrix soaked in thrombin solution can accomplish this mechanical result. Fibrin foam is such a matrix formed from human fibrinogen and thrombin. It effectively controls bleeding from oozing surfaces and from veins, even very large ones. Though not recommended for brisk arterial hemorrhage, it has proved very effective in neurosurgery in controlling hemorrhage from the dura, from tumor beds, from dural sinuses, and from large cerebral veins. In general surgery it has been reported effective, in a small number of patients, in controlling hemorrhage from the cut surface of the liver and kidney, in jaundiced patients, in thoracic operations, and in stopping bleeding in hemophilia.

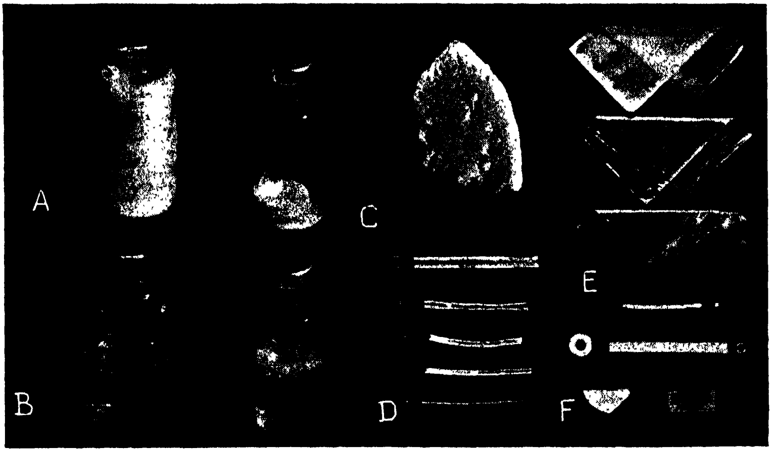


FIG. 106. A, Fibrinogen (left) and thrombin (right) in the dry state. B, Fibrinogen (left) and thrombin (right) in solution. C, Fibrin foam. D, Fibrin tubes. E, Fibrin films. F, Fibrinogen plastics.

In the dry state fibrin foam is a porous material composed of strands of fibers separated by air spaces of macroscopic size. It readily absorbs water, saline, or thrombin solution, and can be used in conjunction with a sulfa drug or with penicillin. The physical properties of the foam, as of the film, can be varied by controlling the conditions of manufacture. Since fibrin foam is prepared completely from human proteins and is rapidly absorbed with minimal tissue reaction, it is left in place, thus preventing recurrence of bleeding. It is easily handled in the operating room, and, like serum albumin, is immediately available for use in emergencies.

The fractionation of plasma, having made available fibrinogen, thrombin, and the fibrinolytic enzyme in physical states such that they can be employed in a wide variety of products, opens a new chapter, not only in the large number of surgical uses which are emerging but also in our fundamental understanding of the relation between the properties of the protein molecules and of the structures that can be prepared from them under controlled conditions.

In 1899 the late Sir William Hardy left the field of histology and developed colloid chemistry in the belief that fundamental advances in our knowledge of the structure of living tissue would have to await advances of our knowledge of the molecules of which it is composed. The time would appear to be at hand when the knowledge of the proteins which Hardy next studied had reached a point where their study in the solid state could form the basis of a new chemical morphology.

THE BLOOD GROUPING GLOBULINS (ISOAGGLUTININS) OF HUMAN PLASMA

The dangers of transfusion with human blood have largely been eliminated by our increasing knowledge of specific differences in the red blood cells of different individuals. Landsteiner reported in 1900 the presence of agglutinogens in red cells and the corresponding isoagglutinins in human sera. Our present classification of blood types and of red cells distinguishes groups, O, A, B, and AB. Other studies have revealed in addition subgroups of type A and AB cells, among them A_2 cells and A_2B cells. The proportion of the individuals in a population belonging to each blood group is approximately as shown in Table V.

The cells of group O blood are not agglutinated by either anti-A or anti-B serum. The blood of such donors is therefore recommended for whole blood transfusion in military medicine, since no reaction normally occurs regardless of the type of the recipient.

The sera of O type blood as well as of A and B type blood contain isoagglutinins, however. If the plasma of O blood is mixed with the proper proportion of B cells, the anti-B isoagglutinins are removed, leaving the plasma rich in anti-A isoagglutinins. The plasma, or serum, of group B bloods is of course also rich in anti-A isoagglutinins, the plasma or serum of type A blood in anti-B isoagglutinins.

Despite the presence of these isoagglutinins in serum or plasma individuals of all blood groups may be transfused with a pooled serum or plasma. The plasma is rendered safe partly by neutralization within the pool. The specific A and B substances have recently been injected in more or less pure state and shown to increase the titer of the recipient's plasma isoagglutinins.

TABLE V

Distribution of Population among Blood Groups

(Computed from various sources for the population of the U. S.)

Group of red blood cells	*Approximate frequency per cent*	*Blood grouping globulins present in plasma**
O	45	Anti-A, Anti-A_1, Anti-B
B	10	Anti-A, Anti-A_1
A_1	31	Anti-B
A_2	10	Anti-B
A_1B	2.6	None
A_2B	1.0	None

* Omitting certain rare isoagglutinins.

The isoagglutinins of the plasma are euglobulins, which are extremely insoluble in the absence of salt near their isoelectric points. Readily separated and concentrated over thirtyfold with respect to the group specific serum or plasma from which they are derived, they can be used in the typing of whole blood. Indeed they may be prepared with such high titers and avidity for red cells of the opposite type that gross clumping observable to the naked eye is brought about in a matter of seconds. The blood grouping globulins are also being prepared from Fraction II + III of group-specific plasma derived from the blood of Red Cross donors and are beginning to be used by the Navy for typing when whole blood transfusion is indicated.

THE IMMUNE PROPERTIES OF THE -GLOBULIN ANTIBODIES OF NORMAL AND CONVALESCENT PLASMA

Many of the antibodies to the variety of infectious diseases to which man has been exposed are γ-globulins. These substances have in the last few years been concentrated and extensively studied not only chemically but also immunologically. The antibodies to many of the common infectious agents to which the population contributing the blood had been exposed have been found. Indeed the large number of blood donors in the Red Cross Blood Donor Program has made statistical deduction valid and demonstrated that a population shows heightened immunity following an epidemic of a special disease, such as the recent influenza epidemic.

The possibility of utilizing the γ-globulin antibodies derived from pooled normal plasma in the prophylaxis and therapy of any infectious disease depends upon two factors: first, the concentration of specific antibody in the blood of the adult population, which in turn is dependent on previous exposures and on the persistence of these bearers of immunity in the body and their distribution between blood and tissues; and second, the amount of antibody necessary to protect an individual against the infection by passive immunization.

The concentration over the pooled plasma thus far achieved by our fractionation process is approximately twenty-five-fold if we compare plasma with the concentrated solution of immune globulins being made available to the Armed Forces, and through the Red Cross, to public health agencies. As a result titers of certain antibodies, though not necessarily of all, are comparable to those of the corresponding convalescent sera. Therefore, in the prevention or treatment of any disease where the value of convalescent serum has been demonstrated, the effectiveness of a concentrate from pooled normal plasma should be investigated. If convalescent serum, or hyper-immune serum is fractionated, still higher antibody titers may be obtained in the concentrated γ-globulins, and their possible value in special cases must also be considered.

The concentrated antibodies, separated from the pooled plasma of a population, offer a means of characterizing the state of immunity of that population as a permanent public health record. Were such records systematically collected and available, it might prove possible to follow the course of epidemics much as the course of a comet may be followed by the photographic records of the skies recorded by modern observatories. Not infrequently the first faint trail of a comet has escaped detection and the origin and course been understood only because astronomers now map all of the heavens systematically at all times. The systematic collection and preservation as dry white powders of the concentrated antibodies of diverse populations might comparably simplify the analysis of the course of certain epidemics.

TABLE VI

Average Concentration of Various Antibodies in Solutions of Normal Human Serum γ-Globulins (Fraction II)

*Antibody**	*Concentration referred to plasma*
Typhoid agglutinin, H	19
Mumps, complement fixation	20
Influenza A, complement fixation	18
Influenza A, Hirst test	10
Influenza A, mouse protection	23
Diphtheria antitoxin	25
Streptococcus antitoxin	22
γ-Globulin by electrophoresis	25

* Titrations of Enders.

Not all of the antibodies concentrated from human plasma will be of value in the control of disease. However, in the case of measles, where most susceptibles that are exposed are likely to contract the disease, the concentrated antibodies have been employed with great success, either to assure complete protection or by the use of a smaller dose to modify the severity of the disease thus providing a lasting immunity with slight risk of those complications which make measles a serious illness. The record of the use of the γ-globulin antibodies in the prophylaxis of measles is given in Table VII.

In the case of many diseases those bearers of immunity which are the antibodies of the blood stream cannot be made available in amounts large enough, or injected soon enough to prevent the onset of the disease. In others the incidence of the disease in susceptibles is so small that it would be necessary to immunize passively an entire population in order to offer protection to the one in a thousand who might contract the disease. The blood donors for such an undertaking might not always be available whereas blood donors conscious of the contribution that they were making to the control of other diseases could clearly always be counted on in a society that has contributed as has ours to the blood donor program.

TABLE VII

Effect of Serum γ-Globulin Antibodies in Measles Prophylaxis

Season	*Epidemic*	*Number of cases*	*No measles per cent*	*Mild measles per cent*	*Average measles per cent*
1942	Philadelphia* and Baltimore	891	69	28	3
1942–43	Boston†	281	59	37	4
1943–44	Scattered‡	111	41	51	8
1943–44	New York§	814	78	20	2
	Total	2097	70	27	3

* Stokes, Maris, and Gellis.
† Ordman, Jennings, and Janeway.
‡ Stokes, Gellis, Janeway, and others.
§ Rutstein, Greenberg, and Frant.

SUMMARY

The Natural Functions and Therapeutic Uses of Blood. Blood is a complex tissue consisting of a fluid part, the plasma, in which cells of various kinds are suspended. Most important among these are the leukocytes or white cells, and the erythrocytes or red cells.

The plasma also is complex. It is an aqueous solution of salts, carbohydrates, and lipids, the latter, for the most part, present in combination with certain of the diverse protein molecules which are the chief plasma constituents.

The various components of the blood stream have such complex interrelations, on the one hand with the tissues, on the other with the external environment, that we may well stress the few systems where an understanding of the mechanism is emerging.

Transport of Oxygen by the Hemoglobin Concentrated in the Red Blood Cells. The red blood cells have as their prime function the transport to the tissues of oxygen and the removal from the tissues of carbon dioxide and other waste products. The oxygen is transported in combination with the hemoglobin concentrated within the red cell. Other proteins, notably carbonic anhydrase and a complex enzyme system, represent its less copious protein components.

When there is great blood loss from a wound, when a major operation is to be performed, or when the anemia which frequently occurs in convalescence from wounds must be combated, red blood cells should be supplied in amounts adequate to restore the oxygen-carrying capacity of the blood.

Maintenance of Blood Volume by the Proteins of the Plasma, Especially by the Albumin. Plasma, the fluid part of the blood, consists of over 90 per cent water and 7 per cent protein. Of the proteins, approximately 58 per cent is albumin, 14 per cent α-globulin, 13 per cent β-globulin, 11 per cent γ-globulin, and 4 per cent fibrinogen.

Plasma is used in military medicine in combating shock and in the treatment of burns. In burns there is often so large a loss from the body of the diverse components of the plasma that their replacement is necessary. In shock the function of the plasma is largely to restore or maintain the blood volume so that circulation may remain normal. The proteins dissolved in the plasma maintain the volume of the blood by virtue of the osmotic pressure exerted largely by the albumin.

Albumin, the most copious of the plasma proteins, exerts the largest osmotic pressure and thus the largest effect in maintaining blood volume. Its solutions are also the least viscous and for this reason their injection should impose the smallest burden on the heart. Although the smallest of the plasma proteins, albumin does not traverse the blood-vessel walls in appreciable amounts under normal conditions.

Storage and Utilization of the Proteins of the Plasma in Equilibrium with the Tissues. The mechanism of control of the blood volume involves the impermeability of the normal blood vessels to the plasma proteins. If plasma proteins are injected in a normal person in excess of need, certain of them leave the blood stream without appearing in the urine and are presumably utilized as a source of energy or new protein or are stored in the tissues. The injection of concentrated serum albumin into the blood stream of a normal man is followed by a rapid redistribution of protein between the blood stream and the tissues. Further investigations are necessary to determine whether this is regulated by colloid osmotic pressure, by subtle changes in the permeability of the blood vessels, or specifically for each species of plasma protein. Clearly there is interaction between the various plasma proteins in the blood and the tissues in health and disease.

Solution and Transport of the Lipids of the Blood in More or Less Labile Combination with α and β Globulins. Lipids of diverse kinds are dissolved and transported in the blood, largely in combination, more or less labile, with proteins. Fibrinogen, γ-globulin, thrombin, and albumin are poor, but α- and β-globulin fractions are rich in lipid material. The separation and further purification of these fractions should increase our understanding of their chemical nature, physiological function, and uses in therapy.

Regulation of Bodily Function by Hormones and Enzymes of the Blood. The hemoglobin of the red cells, the albumins and lipoproteins of the plasma are present in large amounts. The enzymes, which regulate chemical reactions within the blood, and the hormones, many of which have their origin in the various glands of the body and vary in concentration from time to time, are present in but small amount. Their regulatory functions generally involve other glands or tissues in the body than those in which they originate. The potential value of each in therapy is as a specific concentrate.

Coagulation of the Blood by Conversion of Prothrombin.

to Thrombin and Fibrinogen to Fibrin Clots, Fibrin Films, or Fibrin Foams. The mechanism by which the loss of blood from a wound or cut surface is limited involves a series of proteins. Of these but two are essential to form the blood clot; thrombin, present in the circulating blood as a precursor, prothrombin, and fibrinogen, the long asymmetric molecule which is the structural element of the clot.

Fibrinogen and thrombin are both separated in large amounts as products of plasma fractionation. Fibrinogen can be prepared as a thermo-setting plastic or combined with thrombin in the form of foams or films.

Fibrin films resemble natural membranes in many respects and have been employed as a protective layer for the central nervous system in the repair of dural defects.

Fibrin foam and thrombin has proven a valuable hemostatic agent in neurosurgery, especially in the treatment of head injuries in this war.

Grouping of the Blood by Means of the Isoagglutinins Separated as Euglobulins. The blood of a donor to be used in transfusion must not be agglutinated by or cause agglutination of the blood of the recipient. The plasma contains isoagglutinins, proteins, which when separated from type-specific blood have been concentrated over thirtyfold with respect to the plasma and utilized in the grouping of blood for transfusion.

Immune Properties of the Blood Concentrated with the γ-Globulins and Made Available for Prophylaxis or Modification of Measles. The bearers of the immune properties of the blood are proteins, for the most part γ-globulins. The plasma collected by the American Red Cross that has been fractionated yields albumin, fibrinogen, and thrombin, the blood typing globulins and the γ-globulins as protein concentrates. γ-globulin, as distributed to the Armed Forces and to public health agencies, through the American Red Cross, is concentrated twenty-five-fold with respect to plasma and has been found to contain a large variety of the antibodies to the diseases common to man.

The studies that have been completed thus far have demonstrated the value of this concentrate in the prophylaxis or modification of measles.

Preservation of the Concentrated Antibody Globulins of the Blood as a Public Health Record of the Course of Epidemics. Although the concentrated antibodies separated from the pooled plasma of a population will not have value in the prophylaxis or treatment of many diseases they characterize the state of immunity of that population as a permanent public health record.

The chemical, clinical, and immunological investigations on blood and its derivatives which have been carried forward during this war have resulted in contributions to clinical medicine and surgery, to immunology and epidemiology, to protein chemistry, and to the methodics of the production and control of biological products. They are opening new vistas in our understanding of natural products and natural systems.

ACKNOWLEDGEMENTS

This work was originally supported by grants from the Rockefeller Foundation and from funds of Harvard University. It was aided early in 1941 by grants from the Committee on Medicine of the National Research Council, which included a grant from the American College of Physicians. Since August, 1941, it has been carried out under contract, recommended by the Committee on Medical Research, between the Office of Scientific Research and Development and Harvard University. The insight of my many fellow members in this department has been immeasurably aided by many colleagues—chemical, clinical, immunological, and pathological—members of this or other universities, of the Armed Forces, of committees of the National Research Council, of the American Red Cross, and of commercial laboratories with Navy contracts for plasma fractionation, who labored with us. Indeed, it would be impossible to list here the very large number of individuals and agencies who have collaborated in this work.

For References see p. 319.

REFERENCES

CHAPTER I

"Psychological Aspects of Military Aviation"

by Walter R. Miles

1. Adams, J. C. Comments on aviation medicine for the navy for 1942. *Jour. Aviat. Med.*, *14*, 71–74, 1943.
2. Adjutant General's Office. Tests and measurements in the United States army. *Education* (Boston), *64*, 49–54, 1943.
3. Staff, Personnel Research Section, Classification and Replacement Branch, The Adjutant General's Office. Personnel research in the army. II. The classification system and the place of testing. *Psychol. Bull.*, *40*, 205–211, 1943.
4. —— Personnel research in the army. III. Some factors affecting research in the army. *Psychol. Bull.*, *40*, 271–278, 1943.
5. Adjutant General's Office. Testing as a part of military classification. *Science*, *97*, 473–478, 1943.
6. Anonymous Psychology for the fighting man. *Infantry Jour.* (Washington), 1943. 456 pp.
7. —— Effects of flight: mental and physical aspects. McGraw-Hill, New York, 1943. 125 pp.
8. —— Aviation medicine in the A. E. F. Government Printing Office, Washington, 1920. 332 pp.
9. —— Air service medical. Government Printing Office, Washington, 1919. 446 pp.
10. —— The medical problems of flying. Medical Research Council, Special Report Ser., No. 53, London, 1920. 272 pp.
11. Armstrong, H. G. Principles and practice of aviation medicine. Williams and Wilkins, Baltimore, 1939. 496 pp.
12. Barr, E. O. Flying men and medicine: the effects of flying upon the human body. Funk and Wagnalls, New York, 1943. 269 pp.
13. Baker, W. A. Further studies of learning aptitude in pilots. B: Prediction of ground school aptitude. *Trans. Kans. Acad. Sci.*, *45*, 287–289, 1942.
14. Bauer, L. H. Aviation medicine, a brief history; the physical qualifications for flying: oxygen want and the use of supplementary oxygen. *Ann. Intern. Med.*, *18*, 15–20, 1943.
15. Bois, J. S. A. The morale of the fighting soldier. *Bull. Canad. Psychol. Ass.*, *3*, 17–20, 1943.

16. Bott, E. A. Training research, R. A. F. *Bull. Canad. Psychol. Ass.*, *3*, 31–34, 1943.
17. Carlson, W. A. Nervous stability. *Flight Surg. Topics*, *3*, 207–208, 1939.
18. Carson, L. D. Indoctrination of flying personnel in physiologic effects of high altitude flying and need for and use of oxygen. *Jour. Aviat. Med.*, *13*, 162–169, 1942.
19. Chant, S. N. F. Psychology as applied in the Royal Canadian Air Force. *Bull. Canad. Psychol. Ass.*, *3*, 34–37, 1943.
20. Civil Aeronautics Administration. Patter for elementary flight maneuvers. *Civil Aeronautics Bull.*, No. 31. Government Printing Office, Washington, 1943. 91 pp.
21. —— Fundamentals of secondary flight maneuvers for civilian pilot training (mimeographed). Department of Commerce, Washington, 1942. 24 pp.
22. Clinton, J., Jr., and Thorn, G. W. Studies on commercial air line pilots. *War Med.* (Chicago), *4*, 363–373, 1943.
23. Cook, F., and Scott, J. C. The effective flight surgeon. *Contact* (Pensacola), *3*, 185–190, 1943.
24. Davis, H. Symposium on the special senses in relation to military problems. *Federation Proc.* (Am. Soc. Exp. Biol.), *2*, 107–109, 1943.
25. Delucchi, J. R. Psychological factors in the training of airplane pilots. *Jour. Aviat. Med.*, *14*, 84–87, 1943.
26. —— Selection of military aircraft pilots; our viewpoints about professional selection and psychological examination. *Jour. Aviat. Med.*, *13*, 234–244, 1942.
27. Directorate Personnel Selection. Personnel selection in the Canadian Army. *Bull. Canad. Psychol. Ass.*, *3*, 14–15, 1943.
28. Dunbar, F. Medical aspects of accidents and mistakes in the industrial army and in the armed forces. *War. Med.* (Chicago), *4*, 161–175, 1943.
29. Flaiz, J. W., and Wolf, E. G. The effect of noise on aviation personnel. *Contact* (Pensacola), *3*, 160–167, 1943.
30. Flanagan, J. C. The selection and classification program for aviation cadets (aircrews, bombardiers, pilots, and navigators). *Jour. of Consult. Psychol.*, *6*, *229–239*, 1942.
31. Franzen, R., and Brimhall, D. R. Problems of consistency arising from CAA medical examinations. *Rep. Div. Res. Civ. Aeronaut. Adm.*, No. 1, 1942. 20 pp.
32. Hall, C. S. A course in army personnel techniques. *Psychol. Bull.*, *40*, 285–289, 1943.
33. Hoff, E. C., and Fulton, J. F. A bibliography of aviation medicine. Thomas, Baltimore, 1942. 237 pp.
34. Kelley, D. McG. Classification of personnel in the British army:

preliminary survey of methods. *War Med.* (Chicago), *3,* 386–392, 1943.

35. Kellum, W. E. Recent developments in selection of candidates for aviation training. *Am. Jour. Psychiat., 100,* 80–84, 1943.
36. Langewiesche, W. The three secrets of human flight. *Harper's Magazine,* No. 1091 (April), 503-512, 1944.
37. Ligon, E. M. The choice of a combat officer. *Union Alumni Rev., 32,* No. 6, 1943.
38. Louttit, C. M. Psychologists in the navy. *Psychol. Bull., 40,* 375–376, 1943.
39. Mashburn, N. C. Mashburn automatic serial action apparatus for detecting flying aptitude. *Jour. Aviat. Med., 5,* 155–160, 1934.
40. McComas, H. C. The aviator, New York, 1922. 214 pp.
41. McFarland, R. A. The psychological effects of oxygen deprivation (anoxemia) on human behavior. *Arch. Psychol.* No. 145, 1932. 135 pp.
42. —— Psychological aspects of flying with special reference to problems of selection. *Internat. Clin., 1,* Ser. 5, 17–35, 1942.
43. —— Some problems in aviation medicine. *Bull. New Eng. Med. Cent., 5,* 1–6, 1943.
44. Melton, A. W. The selection of pilots by means of psychomotor tests. *Jour. Aviat. Med., 15,* 116–123, 1944.
45. Miles, W. R. The pursuitmeter. *Jour. Exp. Psychol, 4,* 77–105, 1921.
46. —— Red goggles for producing dark adaptation. *Federation Proc.* (Amer. Soc. Exp. Biol.), *2,* 109–115, 1943.
47. Mitchell, H. D. Aircrew selection. *Am. Jour. Psychiat., 99,* 354–357, 1942.
48. Rees, J. R. Three years of military psychiatry in the United Kingdom. *Brit. Med. Jour.,* pt. 1, 1–6, 1943.
49. Reinartz, E. G. The school of aviation medicine and the war. *Milit. Surg., 92,* 233–242, 1943.
50. Rycroft, B. W. Right vision in the army. *Brit. Med. Jour.,* pt. 2, 576–677, 1942.
51. Shaw, C. C. Aviation medicine. *Proc. U. S. Nav. Inst., 68,* 1264–1278, 1942.
52. Staff of the Psychological Branch, Office of the Air Surgeon, Headquarters Army Air Forces. The aviation psychology program of the army air forces. *Psychol. Bull., 40,* 759–769, 1943.
53. Staff, Psychological Research Unit No. 1. History, organization, and procedures psychological research unit No. 1, army air forces. *Psychol. Bull., 41,* 103–114, 1944.
54. —— Psychological Research Unit No. 2, and Staff, Department of Psychology, Research Section. Research program on psychomotor tests in the army air forces. *Psychol. Bull., 41,* 307–321, 1944.

55. —— Psychological Research Unit No. 3. Organization and research activities psychological research unit No. 3, army air forces. *Psychol. Bull., 41,* 237–245, 1944.
56. Steinberg, D. L., and Wittman, M. L. Etiologic factors in the adjustment of men in the armed forces. *War Med.* (Chicago), *4,* 129–139, 1943.
57. Turk, H. M. Factors in the care of the flyer. *Jour. Aviat. Med., 14,* 300–307, 1943.
58. Tuttle, A. D. Physiological and psychological characteristics of successful pilots. *Milit. Surg., 88,* 227–237, 1941.
59. Walsh, M. N. The importance of the nervous energy reserve in aviation. *Proc. Mayo Clin., 16,* 707–714, 1941.
60. Whittingham, H. E. Medical research and aviation. *Jour. Roy. Nav. Med. Serv., 26,* 15–24, 1940.
61. Wittson, C. L., Harris, H. I., and Hunt, W. A. An evaluation of the brief psychiatric interview. *Jour. Psychol., 16,* 107–114, 1943.

CHAPTER II

"The Physical Structure and Biological Action of Nerve Cells" by Detlev W. Bronk

1. Adrian, E. D., and Bronk, D. W. The discharge of impulses in motor nerve fibers. *Jour. Physiol., 66,* 81, 1928.
2. —— The mechanism of nervous action. University of Pennsylvania Press, 1932.
3. —— Discharges from vestibular receptors in the cat. *Jour. Physiol., 101,* 389, 1943.
4. Bear, R. S., and Schmitt, F. O. Optical properties of the axon sheaths of crustacean nerves. *Jour. Cell. and Comp. Physiol., 9,* 275, 1937.
5. Bert, P. Barometric pressure (1877) Translated by Hitchcock, M. A., and Hitchcock, F. A. College Book Co., 1943.
6. Brink, F., Jr., and Bronk, D. W. Chemical initiation of rhythmic local responses in nerve preceding trains of propagated impulses. *Am. Jour. Physiol., 133,* P222, 1941.
7. Bronk, D. W., and Stella, G. Afferent impulses in the carotid sinus nerve. *Jour. Cell. and Comp. Physiol., 1,* 113, 1932.
8. —— and Ferguson, L. K. The nervous control of intercostal respiration. *Am. Jour. Physiol., 110,* 700, 1935.
9. —— The cellular organization of nervous function. *Trans. Coll. Physicians* (Philadelphia), *6,* 102, 1938.

10. —— The nervous regulation of visceral functions. Chemical advances in medicine. University of Minnesota Press, 1940.
11. Cole, K. S., and Curtis, H. J. Electric impedance of the squid giant axon during activity. *Jour. Gen. Physiol., 22,* 649, 1939.
12. Davies, P. W., and Brink, F., Jr. Microelectrodes for measuring local oxygen tension in animal tissues. *Rev. Sci. Instruments, 13,* 524, 1942.
13. Erlanger, J., and Gasser, H. S. Electrical signs of nervous activity. University of Pennsylvania Press, 1937.
14. Hodgkin, A. L. Evidence for electrical transmission in nerve. *Jour. Physiol., 90,* 183, 1937.
15. Lorente de No, R. Transmission of impulses through cranial motor nuclei. Symposium on the synapse. Charles C. Thomas, 1939.
16. Ocker, W. C., and Crane, C. J. Blind flight. Naylor Press, 1932.
17. Schmitt, F. O. The ultrastructure of protoplasmic constituents. *Physiol. Revs., 19,* 270, 1939.
18. Scott, G. H. Mineral distribution in the cytoplasm. Frontiers in cytochemistry. Jacques Cattell Press, 1943.

CHAPTER III

"Energy and Vision" by Selig Hecht

1. Arago, F. Oeuvres complètes, x. Paris, 1858.
2. Barnes, R. B., and Czerny, M. Lässt sich ein Schroteffeckt der Photonen mit dem Auge beobachten? *Zeitschr. f. Physik, 79,* 436, 1932.
3. Boll, F. Zur Anatomie und Physiologie der Retina. *Monatsber. Berlin Akad.,* 783, 1876.
4. Chariton, J., and Lea, C. A. Some experiments concerning the counting of scintillations produced by alpha particles. *Proc. Roy. Soc.* (London), Ser. A, *122,* pt. 1, 304, 1929.
5. Chase, A. M., and Haig, C. The absorption spectrum of visual purple. *Jour. Gen. Physiol., 21,* 411, 1938.
6. Dartnall, H. J. A., and Goodeve, C. F. Scotopic luminosity curve and the absorption spectrum of visual purple. *Nature, 139,* 409, 1937.
7. Fry, T. C. Probability and its engineering uses, p. 476. Van Nostrand, New York, 1928.
8. Hecht, S. Rods, cones, and the chemical basis of vision. *Physiol. Rev., 17,* 239, 1937.

9. ——The chemistry of visual substances. *Ann. Rev. Biochem., 11,* 465, 1942.
10. ——Haig, C., and Chase, A. M. The influence of light adaptation on subsequent dark adaptation of the eye. *Jour. Gen. Physiol., 20,* 831, 1937.
11. ——and Pickels, E. G. The sedimentation constant of visual purple. *Proc. Nat. Acad. Sci.* (Washington), *24,* 172, 1938.
12. ——and Pirenne, M. H. The sensibility of the nocturnal long-eared owl in the spectrum. *Jour. Gen. Physiol., 23,* 709, 1940.
13. ——Shlaer, S., and Pirenne, M. H. Energy, quanta, and vision. *Jour. Gen. Physiol., 25,* 819, 1942.
14. ——and Williams, R. E. The visibility of monochromatic radiation and the absorption spectrum of visual purple. *Jour. Gen. Physiol., 5,* 1, 1922.
15. Koenig, A. Über den menschlichen Sehpurpur und seine Bedeutung für das Sehen. *Sitzungsber. k. Akad. Wissensch. Berlin,* p. 577, 1894.
16. von Kries, J., and Eyster, J. A. E. Über die zur Erregung des Sehorgans erforderlichen Energiemengen. *Zeitschr. Sinnesphysiol., 41,* 394, 1907.
17. Keuhne, W. Chemische vorgänge in der Netzhaut. In: Handbuch der Physiologie, ed. by L. Hermann, *3,* pt. 1, 235. Vogel, Leipzig, 1879.
18. Langley, S. P. Energy and vision. *Phil. Mag.,* Ser. 5, *27,* 1, 1889.
19. Ludvigh, E., and McCarthy, E. F. Absorption of visible light by the refractive media of the human eye. *Arch. Ophth.* (Chicago), *20,* 37, 1938.
20. Lythgoe, R. J. The absorption spectra of visual purple and of indicator yellow. *Jour. Physiol., 89,* 331, 1937.
21. Østerberg, G. Topography of the layer of rods and cones in the human retina. *Acta Ophth.* (Copenhagen), Suppl. 6, 106 pp., 1935.
22. Piper, H. Über Dunkeladaptation. *Zeitschr. Psychol. u. Physiol. Sinnesorgane, 31,* 161, 1903.
23. Roggenbau, C., and Wetthauer, A. Über die Durchlässigkeit der brechenden Augenmedien für langwelliges Licht nach Untersuchungen am Rindsauge. *Klin. Monatsbl. Augenheilk., 79,* 456, 1927.
24. Schultze, M. Zur Anatomie und Physiologie der Retina. *Arch. mikr. Anat., 2,* 175, 1866.
25. Wald, G. Carotenoids and the visual cycle. *Jour. Gen. Physiol., 19,* 351, 1935.
26. ——Area and visual threshold. *Jour. Gen. Physiol., 21,* 269, 1938a.
27. ——On rhodopsin in solution. *Jour. Gen. Physiol., 21,* 795, 1938b.
28. Weaver, K. S. The visibility of radiation at low intensities. *Jour. Opt. Soc. Am., 27,* 36, 1937.

CHAPTER IV

"Chemical Transmission of Nervous Impulse"

by Otto Loewi

1. Elliott. *Jour. Physiol., 32,* 401, 1905.
2. Dixon. *Medical Magazine,* 1907.
3. Elliott. *Jour. Physiol., 31,* Proc. xx, 1904. 20 pp.
4. Loewi. *Pfluegers Arch., 189,* 239, 1921.
5. Hunt and Taveau. *Brit. Med. Jour.,* 1788, 1906.
6. Loewi and Navratil. *Pfluegers Arch., 214,* 678, 1926.
7. —— and Navratil. *Pfluegers Arch., 214,* 689, 1926.
8. Dale and Dudley. *Jour. Physiol., 68,* 97, 1929.
9. Loewi. *Pfluergers Arch., 237,* 504, 1936.
10. Cannon and Rosenblueth. *Am. Jour. Physiol., 104,* 557, 1933.
11. Engelhart. *Pfluegers Arch., 227,* 220, 1931.
12. Dale. *Jour. Physiol., 81,* Proc. x, 1933. 10 pp.
13. Feldberg and Gaddum. *Jour. Physiol., 81,* 305, 1934.
14. Symposium on the Synapse. *Jour. Neurophysiol., 12,* 361, 1939.
15. Engelhart and Loewi. *Arch. Intern. der Pharmacodyn., 38,* 287, 1930.
16. Quastel, Tennenbaum, and Wheatley. *Biochem. Jour., 30,* 1668, 1936.
17. MacIntosh and Kahlson. *Jour. Physiol., 96,* 277, 1939.
18. Eccles. *Physiol. Rev., 17,* 538, 1937.
19. Brown and Feldberg. *Jour. Physiol., 86,* 290, 1936.
20. Loewi and Navratil. *Pfluegers Arch., 206,* 123, 1924.
21. Navratil. *Pfluegers Arch., 217,* 610, 1927.
22. Feldberg and Vartiainen. *Jour. Physiol., 83,* 103, 1934.
23. Brown and Feldberg. *Jour. Physiol., 88,* 265, 1936.
24. Brown, Dale, and Feldberg. *Jour. Physiol., 82,* 394, 1936.
25. Nachmansohn. *Jour. Neurophysiol., 6,* 397, 1943.
26. Dale, Feldberg, and Vogt. *Jour. Physiol., 86,* 353, 1936.
27. Nachmansohn. *Yale Jour. Biol. and Med., 12,* 565, 1940.
28. Marnay and Nachmansohn. *Jour. Physiol., 89,* 359, 1937; *92,* 37, 1938.
29. Sherrington. *Proc. Roy. Soc.* (London), Ser. B, *97,* 519, 1925.
30. Nachmansohn. *Bull. Soc. de Chem. Biol., 21,* 761, 1939.
31. MacIntosh. *Jour. Physiol., 99,* 436, 1941.
32. Dickshit. *Jour. Physiol., 80,* 409, 1934.
33. Cannon and Haimovici. *Am. Jour. Physiol., 126,* 731, 1939.
34. Buelbring and Burn. *Jour. Physiol., 100,* 337, 1941.
35. Miller, Stavraky, and Woonton. *Jour. Neurophysiol., 3,* 131, 1940.

CHAPTER V

"The Mathematical Nature of Physical Theories"

by George D. Birkhoff

1. Logic: the theory of inquiry, p. 35. New York, 1938.
2. See, for instance, Birkhoff, Garrett, Lattice theory. New York, 1942.
3. In our elementary text, Basic geometry (Chicago, 1941), Professor Ralph Beatley and I have inserted a set of "laws of number."
4. See Stone, M. H., The theory of representations for Boolean algebras. *Trans. Am. Math. Soc., 40,* 1936. Unfortunately no elementary exposition of the central results is available.
5. This set of postulates for the "time series" appears to be novel in the postulate IV of the isomorphism of all (open) intervals. The postulate V, of a type due to Souslin, is found to be necessary. Since my Sigma Xi lectures were given, a complete proof has been obtained that my conjecture is valid. See an article Los Continuos Lineales Homogeneos de George D. Birkhoff by Roberto Vásquez and Francisco Zubieta, *Boletin de la Sociedad Matemática Mexicana, 1,* No. 2 (January, 1944).
6. The principle of sufficient reason. Rice Institute Pamphlet, January, 1941.
7. See Fee, Jerome, Maupertuis and the principle of least action. *Am. Scientist,* April, 1942, and the *Scientific Monthly,* June, 1941, for a discussion of this principle.
8. See my paper presented on February 20, 1942, at the occasion of the Pan-American Astro-Physical Congress held at Puebla, Mexico, in connection with the inauguration of the Astronomical Observatory at Tonanzintla, under the title El Concepto de Tiempo y la Gravitación. It will be published in the *Proceedings* of the Congress. Also see my article, Newtonian and other forms of gravitational theory, *Scientific Monthly,* January and February, 1944, and a brief technical account in the *Revista de Ciencias* of the University of San Marcos, Lima, No. 441, 1942. A more adequate technical account is to be found in my note, Matter, electricity and gravitation in flat space-time, published in August, 1943, in *Proc. Nat. Acad. Sci.* At the present time Professors Alberto Barajas, Carlos Graef, and Manuel Sandoval Vallarta of the National University of Mexico, and I are developing the theory and its applications further. Our results will appear shortly.
9. On relativity and convection, Mathematical and physical papers, *1,* 1929.
10. A theory of matter and electricity, and The hydrogen atom and the Balmer formula, *Proc. Nat. Acad. Sci.,* March, 1927.

11. If there are n bodies in the system, the values of the masses would have to be thought of as increased in a certain ratio n(n—1)/2.
12. Quantum mechanics, p. 7. Rice Institute, 1930.
13. *Ibid.*, p. 45.
14. That is, uses multi-dimensional symbols beyond our grasp.

CHAPTER VI

"The Magnetic Approach to the Absolute Zero of Temperature" by Peter Debye

1. Keesom, W. H. *Proc. Kon. Akad. Amsterdam, 35,* 136, 1932.
2. Debye, P. *Ann. d. Phys., 81,* 1154, 1926.
   Giauque, W. F. *Jour. Am. Chem. Soc., 49,* 1864, 1927.
3. Curie, P. *Jour. de Phys., 4,* 197, 1895.
4. Langevin, P. *Chem. Rev., 139,* 1204, 1904; *Jour. de Phys., 4,* 678, 1905; *Ann. de Chimie et Phys., 5,* 70, 1905.
5. Woltjer, H. R., and Kamerlingh-Onnes, H. *Leiden Comm.,* 167c, 1923; *Versl. Kon. Akad. Amsterdam, 32,* 772, 1923.
6. Giauque, W. F., and MacDougall, D. P. *Phys. Rev., 43,* 768, 1933; *44,* 235, 1933.
7. de Haas, W. J., Wiersma, E. C., and Kramers, H. A. *Naturwiss., 21,* 467, 1933; *Nature, 131,* 719, 1933; *132,* 372, 1933; *Phys., 13,* 175, 1933.
   Kramers, H. A. *Leipziger Vorträge,* 1933.
   de Haas, W. J., and Wiersma, E. C. *Leiden Comm.,* Suppl. 81C.
8. de Haas, W. J., and Wiersma, E. C. Comprehensive report in *Rapports du septième Congrès international du Froid,* 1936. Many separate articles in *Leiden Comm.*
9. van Vleck, J. H. *Jour. Chem. Phys., 5,* 320, 1937.
   Hebb, M. H., and Purcell, E. M. *Jour. Chem. Phys., 5,* 338, 1937.
   Debye, P. *Ann. d. Phys., 32,* 85, 1938.
10. Lipson, H., and Beevers, C. A. *Proc. Roy. Soc.* (London), *148,* 664, 1935.
11. Kürti, N., Lainé, P., Rollin, B. V., and Simon, F. *Chem. Rev., 204,* 1576, 1936.
12. Casimir, H. B. G., de Haas, W. J., and de Klerk, D. *Phys., 6,* 241, 1939; *Leiden Comm.,* 256a.
13. Waller, I. *Zeitschr. f. Physik, 104,* 132, 1936.
14. Keesom, W. H. *Phys. Zeitschr., 35,* 9, 1934.
15. Gorter, C. J., and Brons, F. *Phys., 4,* 579, 667, 1937; *Nature, 141,* 369, 1938.
16. ——Teunissen, P., and Brons, F. *Phys., 5,* 657, 1938.
17. Kürti, N., Lainé, P., and Simon, F. *Chem. Rev., 204,* 754, 1937.

Books:

Casimir, H. B. G. Magnetism and very low temperatures. Cambridge University Press, 1940.

Ruhemann, M., and B. Low temperature physics. Cambridge University Press, 1937.

van Lammeren, J. A. Technik der tiefen temperaturen. J. Springer, Berlin, 1941.

van Vleck, J. H. The theory of electric and magnetic susceptibilities. Oxford University Press, 1932.

CHAPTER VII

"The Drift Toward Equilibrium" by Henry Eyring

1. Condon, E. U. *Phys. Rev. 54,* 937, 1938.
   Eyring, H., Walter, John, and Kimball, G. E. Quantum chemistry. John Wiley and Sons, 1944.
2. —— *Jour. Chem. Phys., 3,* 107, 1935.
3. Pelzer, H., and Wigner, E. *Zeitschr. f. physik. Chem.,* B*15,* 445, 1932.
   Wigner, E. *Zeitschr. f. physik. Chem.,* B*19,* 203, 1932.
4. Eyring, H., and Polanyi, M. *Zeitschr. f. physik. Chem.,* B*12, 279,* 1931.
5. London, F. *Zeitschr. f. Elektrochem., 35,* 552, 1929.
6. Marcelin, R. *Ann. Physik, 3,* 120, 1915.
7. Evans, M. G., and Polanyi, M. *Trans. Faraday Soc., 31,* 875, 1935.
8. Wynne-Jones, W. F. K., and Eyring, H. *Jour. Chem. Phys., 3,* 492, 1935.
9. Glasstone, S., Laidler, K. J., Eyring, H. The theory of rate processes. McGraw-Hill, 1941.
10. Fowler, R. H., and Guggenheim, E. A. Statistical thermodynamics. Cambridge University Press, 1939.
11. Hirschfelder, J. O., Eyring, H., and Topley, B. *Jour. Chem. Phys., 4,* 170, 1936.
12. —— and Wigner, E. *Jour. Chem. Phys., 7,* 616, 1939.
13. Eyring, H. *Jour. Chem. Phys., 4,* 283, 1936.
14. Batschinski, A. J. *Zeitschr. f. physik. Chem., 84,* 643, 1913.
15. Walter, J., and Eyring, H. *Jour. Chem. Phys., 9,* 393, 1941.
16. Powell, R. E., Roseveare, W. E., and Eyring, H. *Ind. and Eng. Chem., 33,* 430, 1941.
17. Hughes, E. D., and Ingold, C. K. *Trans. Faraday Soc., 37,* 657, 1941.

Baughan, E. C., Evans, M. G., and Polanyi, M. *Trans. Faraday Soc., 37,* 377, 1941.

18. Eyring, H. *Zeitschr. f. physik. Chem., 7,* 244, 1930.
    Kassel, L. S. Kinetics of homogeneous reactions, p. 72. The Chemical Catalog Co., 1932.
19. Hirschfelder, J. O. *Jour. Chem. Phys., 9,* 645, 1941.
    Daniels, F. *Ind. and Eng. Chem., 35,* 504, 1943.
20. Pauling, L. Nature of the chemical bond. Cornell University Press, 1940.
21. Altar, W., and Eyring, H. *Jour. Chem. Phys., 4,* 661, 1936.
22. Morse, P. M. *Phys. Rev., 34,* 57, 1929.
23. Heitler, W., and London, F. *Zeitschr. f. Physik, 44,* 455, 1927.
    Sugiura, Y. *Zeitschr. f. Physik., 45,* 484, 1927.
24. Rosen, N., and Ikehara, S. *Phys. Rev., 43,* 5, 1933.
25. Magee, J. L. *Jour. Chem. Phys., 8,* 677, 1940.
26. Harvey, E. N. Living light. Princeton University Press, 1940; *Ann. Rev. Biochem., 10,* 531, 1941.
27. Anderson, R. S. *Jour. Cell. and Comp. Physiol., 8,* 261, 1936.
    van Schowenburg, K. L. Respiration and light emission in bacteria. Thesis. Delft, 1938.
    Johnson, F. H. *Enzymologia, 7,* 72, 1939.
    —— van Schowenburg, K. L., and van der Burg, A. *Enzymologia, 7,* 195, 1939.
    Chance, B., Harvey, E. N., Johnson, F. H., and Millikan, G. *Jour. Cell. and Comp. Physiol., 15,* 195, 1940.
28. Brown, D. E., Johnson, F. H., and Marsland, D. A. *Jour. Cell. and Comp. Physiol., 20,* 151, 1942.
    Eyring, H., and Magee, J. L. *Jour. Cell. and Comp. Physiol., 20,* 169, 1942.
29. Anson, M. L., and Mirsky, A. E. *Jour. Gen. Physiol., 17,* 393, 1923.
    Stearn, A. E., and Eyring, H. *Jour. Chem. Phys., 5,* 113, 1937.
    Eyring, H., and Stearn, A. E. *Chem. Rev., 25,* 253, 1939.
30. Johnson, F. H., Eyring, H., and Williams, R. W. *Jour. Cell. and Comp. Physiol., 20,* 247, 1942.
31. McElroy, William D. *Jour. Cell. and Comp. Physiol., 21,* 95, 1943.
32. Johnson, F. H., Eyring, H., and Kearns, W. *Arch. Biochem., 3,* 1, 1943.
33. Steinhart, J. *Kgl. Danske Videnskab Selskab, Math.-fys. Medd., 14,* 11, 1937.
34. La Mer, V. K. *Science, 86,* 614, 1937.
    Eyring, H., and Stearn, A. E. *Chem. Rev., 25,* 253, 1939.
    Sizer, I. W. Advances in enzymology, *3,* 44. Interscience Press, 1943.

CHAPTER VIII

"Streams of Atoms" by I. I. Rabi

Fraser, R. G. S. Molecular rays. Macmillan Co., 1931.
Rabi, I. I. *Phys. Rev., 51,* 652, 1937.
—— Millman, Kusch, and Zacharias. *Phys. Rev., 55,* 526, 1939.
Kellogg, Rabi, Ramsey, and Zacharias. *Phys. Rev., 56,* 728, 1939.
—— Rabi, Ramsey, and Zacharias. *Phys. Rev., 57,* 677, 1940.
Kusch, Millman, and Rabi. *Phys. Rev., 57,* 765, 1940.

CHAPTER IX

"Adventures in Vacuum Chemistry" by K. C. D. Hickman

1. Burch, C. R. *Nature, 122,* 729, 1928.
2. Embree, N. D. *Chem. Rev., 29,* 317, 1941.
3. Jewell, W., Mead, T. H., and Phipps, J. W. *Jour. Soc. Chem. Ind., 58,* 56, 1939.
4. Burch, C. R. *Proc. Roy. Soc.* (London), *A123,* 271, 1929.
   —— et al. British patent 303,078, December 21, 1928; reissued as U. S. patent 1,955,321, April 17, 1934.
5. Hickman, K. C. D. *Ind. and Eng. Chem., 29,* 968, 1937.
6. —— *Chem. Rev., 34,* 51, 1944.
7. Burch, C. R., and Van Dijck, W. J. D. *Jour. Soc. Chem. Ind., 58,* 39, 1939.
8. Fawcett, E. W. *Jour. Soc. Chem. Ind., 58,* 43, 1939.
9. Hickman, K. C. D. *Jour. Franklin Inst., 221,* 383, 1936.
10. Wilson, R. E. U. S. patent 1,905,201, April 25, 1933.
11. Embree, N. D. *Ind. and Eng. Chem., 29,* 975, 1937.
12. Hickman, K. C. D., U. S. patent 2,113,302, April 5, 1938.
13. Bills, C. E., Massengale, O. N., Hickman, K. C. D., and Gray, E. Le B. *Jour. Biol. Chem., 126,* 241, 1938.
14. Robeson, C. D., and Baxter, J. G. *Jour. Am. Chem. Soc., 65,* 910, 1943.
15. Lederer, E., and Rosanova, V. *Biokhimiya, 2,* 293, 1937.
16. Gray, E. Le B. *Jour. Biol. Chem., 131,* 317, 1939.
17. Embree, N. D., and Shantz, E. M. *Jour. Am. Chem. Soc., 65,* 910, 1943.
18. Gray, E. Le B., Hickman, K. C. D., and Brown, E. F. *Jour. Nutr., 19,* 39, 1940.
19. Davies, A. W., and Moore, T. *Nature, 147,* 794, 1941.

20. Hickman, K. C. D., Kaley, M. W., and Harris, P. L. *Jour. Biol. Chem., 152,* 303, 313, 321, 1944.
21. Dam, H., and Lewis, L. *Biochem. Jour., 31,* 17, 1937.
    —— and Schönheyder, F. *Nord. med. Tid., 12,* 1097, 1936.
22. Almquist, H. J. *Jour. Biol. Chem., 115,* 589, 1936.
    Breusch, F. *Zeitschr. physiol. Chem., 227,* 242, 1934.
23. Bancroft, F. E. British patent 428,719, May 15, 1935.
24. Burrows, G. *Jour. Soc. Chem. Ind., 58,* 50, 1939.
25. Hecker, J. C. U. S. patent 2,180,052, November 14, 1939.
26. Detwiler, S. B., Jr. Oil and Soap, *16,* 2, 1939; *17,* 241, 1940.
27. —— Abstract of articles and patents on molecular or short-path distillation. U. S. Regional Soybean Industrial Products Laboratory, Urbana, Illinois, 1941.
28. Todd, S. S. *Oil and Soap, 20,* 205, 1943.
29. Waterman, H. I., and Oosterhof, D. *Rec. Trav. Chim., 52,* 895, 1933.
30. Baxter, J. G., and Robeson, C. D. *Jour. Am. Chem. Soc., 64,* 2407, 1942.
    —— and Robeson, C. D. *Jour. Am. Chem. Soc., 64,* 2411, 1942.
    —— Robeson, C. D., Taylor, J. D., and Lehman, R. W. *Jour. Am. Chem. Soc., 65,* 918, 1943.
    Joffe, M., and Harris, P. L. *Jour. Am. Chem. Soc., 65,* 925, 1943.

CHAPTER X

"Present Status of the Vitamin B Complex"

by C. A. Elvehjem

1. Williams, R. R., and Spies, T. D. Vitamin B_1 and its use in medicine. Macmillan, New York, 1938.
2. Hogness, T. R. A symposium on respiratory enzymes, p. 134. The University of Wisconsin Press, 1942.
3. Snell, E. E., and Strong, F. M. *Ind. and Eng. Chem.* (Anal. ed.), *11,* 346, 1939.
4. Elvehjem, C. A. *Physiol. Rev., 20,* 249, 1940.
5. Harris, S. A., and Folkers, Karl. *Jour. Am. Chem. Soc., 61,* 245. 1939.
6. Woolley, D. W., Waisman, H. A., and Elvehjem, C. A. *Jour. Am. Chem. Soc., 61,* 977, 1939.
7. Jukes, T. H. *Jour. Am. Chem. Soc., 61,* 975, 1939.
8. Williams, R. J., Lyman, C. M., Goodyear. G. H., Truesdail, J. H., and Haladay, D. *Jour. Am. Chem. Soc., 55,* 2912, 1933.
9. Stiller, E. T., Harris, S. A., Finkelstein, Jacob, Keresztesy, J. C., and Folkers, Karl. *Jour. Am. Chem. Soc., 62,* 1785. 1940.

10. McHenry, E. W. Biological symposia, ed. by H. B. Lewis, *5,* 177. Jaques Cattell Press, Lancaster, Penn., 1941.
11. du Vigneaud, Vincent. Biological symposia, ed. by H. B. Lewis, *5,* 234. Jaques Cattell Press, Lancaster, Penn., 1941.
12. Jukes, T. H. *Jour. Biol. Chem., 134,* 789, 1940.
13. Richardson, L. R., Hogan, A. G., Long, B., and Itschner, K. I. *Proc. Soc. Exp. Biol. and Med., 46,* 530, 1941.
14. Jukes, T. H. *Proc. Soc. Exp. Biol. and Med., 45,* 625, 1940.
15. Woolley, D. W. *Proc. Soc. Exp. Biol. and Med., 43,* 352, 1940.
16. Dimick, M. K., and Lepp, A. *Jour. Nutr., 20,* 413, 1940.
17. Unna, K. *Jour. Nutr., 20,* 565, 1940.
18. Frost, D. V., Moore, R. C., and Dann, F. P. *Proc. Soc. Exp. Biol. and Med., 46,* 507, 1941.
19. Henderson, L. M., McIntire, J. M., Waisman, H. A., and Elvehjem, C. A. *Jour. Nutr., 23,* 47, 1942.
20. Theiler, A., Green, H. H., and Viljoen, P. R. *3d and 4th Reports,* Director Veterinary Res., Dept. Agr., Union S. Africa, 1915.
21. Wegner, M. I., Booth, A. N., Elvehjem, C. A., and Hart, E. B. *Proc. Soc. Exp. Biol. and Med., 47,* 90, 1941.
22. Marshall, E. K., Jr., Bratton, A. C., White, H. J., and Litchfield, J. T., Jr. *Bull. Johns Hopkins Hosp., 67,* 163, 1940.
23. Black, S., McKibbin, J. M., and Elvehjem, C. A. *Proc. Soc. Exp. Biol. and Med., 47,* 308, 1941.
24. Mackenzie, J. B., Mackenzie, C. G., and McCollum, E. V. *Science, 94,* 518, 1941.
25. Black, S., Overman, R. S., Elvehjem, C. A., and Link, K. P. *Jour. Biol. Chem., 145,* 137, 1942.
26. Hutchins, B. L., Bohonos, N., and Peterson, W. H. *Jour. Biol. Chem., 141,* 521, 1941.
27. Welch, A. D. *Federation Proc.* (Am. Soc. Exp. Biol.), *1,* pt. 2, 171, 1942.
28. Nielson, E., and Elvehjem, C. A. *Jour. Biol. Chem., 145,* 713, 1942.
29. Poth, E. I., Knotts, F. L., Lee, J. T., and Inui, F. *Arch. Surg., 44,* 187, 1942.
30. Du Vigneaud, Vincent, Melville, D. B., György, P., and Rose, C. S. *Science, 92,* 62, 1940.
31. Nielsen, E., Shull, G. M., and Peterson, W. H. *Jour. Nutr., 24,* 523, 1942.
32. Daft, F. S., Ashburn, L. L., and Sebrell, W. H. *Science, 96,* 321, 1942.
33. Martin, G. J. *Proc. Soc. Exp. Biol. and Med., 51,* 252, 1942.
34. Hegsted, D. M., Oleson, J. J., Mills, R. C., Elvehjem, C. A., and Hart, E. B. *Jour. Nutr., 20,* 599, 1940.
35. Norris, E. R., and Hauschildt, J. *Science, 92,* 316, 1940.

36. Woolley, D. W. *Jour. Biol. Chem., 136,* 113, 1940.
37. —— *Jour. Exp. Med., 75,* 277, 1942.
38. Stokstad, E. L. R., and Manning, P. D. V. *Jour. Biol. Chem., 125,* 687, 1938.
39. Mills, R. C., Briggs, G. M., Jr., Elvehjem, C. A., and Hart, E. B. *Proc. Soc. Exp. Biol. and Med., 49,* 186, 1942.
40. Rubbo, S. D., and Gillespie, J. M. *Nature, 146,* 838, 1940.
41. Ansbacher, S. *Science, 93,* 164, 1941.
42. Sieve, B. F. *Science, 94,* 257, 1941.
43. Sober, H. A., Mannering, G. J., Cannon, M. D., Elvehjem, C. A., and Hart, E. B. *Jour. Nutr., 24,* 503, 1942.
44. Woolley, D. W. *Jour. Biol. Chem., 143,* 679, 1942.
45. Hogan, A. G., and Hamilton, J. W. *Jour. Nutr., 23,* 533, 1942.
46. Mannering, G. J. Unpublished.
47. Sydenstricker, V. P., Singal, S. A., Briggs, A. P., De Vaugh, N. M., and Isbell, H. *Jour. Am. Med. Ass., 118,* 1199, 1942.
48. Oppel, T. W. *Am. Jour. Med Sci., 204,* 856, 1942.
49. Hegsted, D. M., Mills, R. C., Briggs, G. M., Jr., Elvehjem, C. A., and Hart, E. B. *Jour. Nutr., 23,* 175, 1942.
50. Landy, M., and Dicken, D. M. *Proc. Soc. Exp. Biol. and Med., 46,* 449, 1941.
51. Folch, J., and Woolley, D. W. *Jour. Biol. Chem., 142,* 963, 1942.
52. Mitchell, H. K., and Isbell, E. R. *The University of Texas Publication,* No. 4237, p. 125, 1942.
53. Day, P. L., Langston, W. C., and Shukers, C. J. *Jour. Nutr., 9,* 637, 1935.
    —— Langston, W. C., Darby, W. J., Wahlen, J. G., and Mims, V. *Jour. Exp. Med., 66,* 579, 1937.
54. Saslow, S., Schwab, J. L., Woolpert, O. C., and Wilson, H. E. *Proc. Soc. Exp. Biol. and Med., 51,* 391, 1942.
55. Spicer, S. S., Daft, F. S., Sebrell, W. H., and Ashburn, L. L. *Pub. Health Rep., 57,* 1559, 1942.
56. Ashburn, L. L., Raft, F. S., Endicott, K. M., and Sebrell, W. H. *Pub. Health Rep., 57,* 1883, 1942.
57. Long, P. H. *Jour. Am. Med. Ass., 121,* 1943.

CHAPTER XI

"Blood and Blood Derivatives" by Edwin J. Cohn

A. Blood

Our present knowledge of the respiratory function of the blood is based so largely on the work of the Schools of Physiology in Cambridge, England (J. Barcroft, The respiratory function of the blood, Cambridge University Press, Cambridge, 1914), and in Cambridge, Massachusetts (L. J. Henderson, Blood, a study in general physiology, Yale University Press, New Haven, 1928) that reference is made to these volumes for the earlier literature. Our knowledge of carbonic anhydrase also comes from the Cambridge School (F. J. W. Roughton, *Phys. Rev., 15,* 241, 1935), and from the Connaught Laboratories where it has recently been crystallized by Scott (D. A. Scott and A. M. Fisher, *Jour. Biol. Chem., 144,* 371, 1942). "The mammalian red cell and the properties of haemolytic systems" are discussed by E. Ponder (Gebruder Borntraeger, Berlin, 1934, and more recently in *Cold Spring Harbor Symposium on Quantitative Biology, 8,* 1940, and in a review by Higgins (*Ann. Rev. Physiol., 3,* 283, 1941). A great deal of new knowledge has been gained in this war but has not yet been made available.

"A plan for collection, transportation and administration of whole blood and of plasma in warfare" (DeGowin and Hardin, War medicine, *1,* 326, 1941) reviews the earlier experience on blood transfusion.

The technique of blood preservation and of rotating blood banks has recently been summarized in a bulletin prepared by the Medical Division of the Office of Civilian Defense (OCD Publication 2220, March, 1944).

B. Dry Plasma

D'Arsenval and Bordas (*Compt. rend. des Séances de l'Acad. des Sci., 142,* 1058, May, 1906) introduced the use of low temperature for condensation of water vapor with a high vacuum for drying of easily alterable substances such as starches and albuminoid substances. In 1934 Elser (U. S. Patent 1970956, 1934) and in 1935 Elser, Thomas, and Steffen (*Jour. Immunol., 28,* 433, 1935) described a process and apparatus for desiccating liquid and semi-solid biological substances by means of high vacuum low temperature condensation.

Since 1935, Flosdorf and Mudd (*Jour. Immunol., 29,* 389, 1935; *ibid., 34,* 469, 1938), and Flosdorf, Stokes, and Mudd (*Jour. Am. Med. Ass., 115,* 1095, 1940), have published a series of papers on drying from the frozen state. Important contributions have since been made by Greaves and Adair (*Jour. Hyg., 39,* 413, 1939), Hill and Pfeiffer (*Ann. Intern. Med., 14,* 201, 1940), Strumia and McGraw (*Jour. Lab. and Clin. Med., 28,* 1140, 1943), and Strumia, McGraw, and Reichel (*Am. Jour. Clin. Path., 11,* 480, 1941). The use of dry plasma had been extensively investigated by Strumia, Wagner, and Monaghan (*Ann. Surg., 111,* 623, 1940; *Jour. Am. Med. Ass., 114,* 1337, 1940), Strumia and McGraw (*Jour. Am. Med. Ass., 116,* 2378, 1941), and others before this war.

C. The Physical Properties of Plasma Components

The molecular weights and shapes of the plasma proteins are discussed by Svedberg and Petersen (The ultracentrifuge, Oxford University Press, Oxford, 1940), Oncley (*Jour. Phys. Chem., 44,* 1103, 1940; *Chem. Rev., 30,* 433, 1942), Tiselius (*Biochem. Jour., 31,* 313, 1464, 1937); *Trans. Faraday Soc., 33,* 524, 1937), Cohn (*Chem. Rev., 28,* 395, 1941; *Trans. and Stud., Coll. Physicians* [Philadelphia], Ser. 4, *10,* 149, 1942) and most recently by Cohn, Oncley, Strong, Hughes, and Armstrong (*Jour. Clin. Inv., 23,* 417, 1944).

Electrophoretic analyses of the plasma proteins in health and disease are given by Tiselius (*Trans. Faraday Soc., 33,* 524, 1937; *Biochem. Jour., 31,* 1464, 1937) and Tiselius and Kabat (*Jour. Exp. Med., 69,* 119, 1939), Svensson (*Kolloid-Zeitschr., 87,* 181, 1939), Longsworth, Shedlovsky, and MacInnes (*Jour. Exp. Med.,* 70, 399, 1939), MacInnes and Longsworth (*Science, 89,* 438, 1939), Longsworth and MacInnes (*Chem. Rev., 24,* 271, 1939; *Jour. Exp. Med., 71,* 77, 1940), Luetscher (*Jour. Am. Chem. Soc., 61,* 2888, 1939; *Jour. Clin. Inv., 19,* 313, 1940; and *Jour. Clin. Inv., 20,* 99, 1941).

D. The Fractionation of Plasma

During the last century there have been innumerable investigations on the separation of one or another of the components of plasma. Two classical papers, both of which appeared in 1905 are by Hardy (*Jour. Physiol., 33,* 251, 1905) and Mellanby (*Jour. Physiol., 33,* 338, 1905). Then came some important contributions from Chick and Martin (*Jour. Physiol., 40,* 404, 1910; *43,* 1, 1911; *45,* 61, 261, 1912–13), and from Sorensen ("Proteins," The Fleischmann Laboratories, New York, 1925). Much of the earlier literature on this subject not specifically referred to in this communication is considered by Cohn (*Physiol. Rev., 5,* 349, 1925; *Ann. Rev. Biochem., 4,* 93, 1935; Harvey lectures, No. 124, 1938–39), Tiselius (*Nova Acta Soc. Sci. Upsal., 4,* 7, 1930), Sörensen and Haugaard (*Compt. rend. trav. lab. Carlsberg, 19,* No. 12, 1933), Butler and Montgomery (*Jour. Biol. Chem., 99,* 173, 1932), Butler, Blatt, and Southgate (*Jour. Biol. Chem., 109,* 755, 1935), Kydd (*Jour. Biol. Chem., 107,* 747, 1934), Hewitt (*Biochem. Jour., 30,* 2229, 1936), Stenhagen (*Biochem. Jour., 32,* 714, 1938), Philpot (*Nature, 141,* 283, 1938), Green (*Jour. Am. Chem. Soc., 60,* 1108, 1938), Fahey and Green (*Jour. Am. Chem. Soc., 60,* 3039, 1938), Longsworth (*Jour. Am. Chem. Soc., 61,* 529, 1939), Kekwick (*Biochem. Jour., 33,* 1122, 1939), Blix (*Zeitschr. ges. exp. Med., 105,* 595, 1939), Svedberg and Pedersen (The ultracentrifuge, Oxford University Press, London, 1940). More recent references describing the fractionation of plasma to yield its diverse components as products for clinical appraisal are Cohn, McMeekin, Oncley, Newell, and Hughes (*Jour. Am. Chem. Soc.,* 62, 3386, 1940), McMeekin (*Jour. Am. Chem. Soc., 62,* 3393, 1940), Cohn, Luetscher, Oncley, Armstrong, and Davis (*Jour. Am. Chem. Soc., 62,* 3396, 1940), Cohn (*Chem. Rev., 28,* 395, 1941; *Trans. and Stud., Coll. Physicians* [Philadelphia], *10,* 149, 1942; *Proc. Am. Philos. Soc., 88,* 159, 1944), and Cohn, Oncley, Strong, Hughes, and Armstrong (*Jour. Clin. Inv., 23,* 417, 1944).

E. Serum Albumin

The chemical properties of albumin are also discussed by Cohn (*Chem. Rev., 28,* 395, 1941) and by Cohn, Oncley, Strong, Armstrong, Ferry, and Hughes (in Mudd, S., and Talhimer, W., Blood substitutes and blood transfusion, Charles C. Thomas, Springfield, 1942, chap. xx), its osmotic pressure by Scatchard, Batchelder, and Brown (*Jour. Clin. Inv., 23,* 458, 1944), and its clinical use by Janeway, Gibson, Woodruff, Heyl, Bailey, and Newhouser (*Jour. Clin. Inv., 23,* 465, 1944), Janeway (*Jour. Am. Med., Ass., 126,* 674, 1944), Newhouser and Lozner (*U. S. Nav. Med. Bull., 40,* 796, 1942), Stead and Ebert (in Mudd, S., and Thalhimer, W., Blood substitutes and blood transfusion, p. 185), Heyl, Gibson, and Janeway (*Jour. Clin. Inv., 22,* 763, 1943), Woodruff and Gibson (*U. S. Nav. Med. Bull., 40,* 791, 1942), Kendrick, Reichel, and McGraw (*Army Med. Bull., 68,* 107, 1943), Janeway (in Mudd, S., and Thalhimer, W., Blood substitutes and blood transfusion, chap. xxi), Luetscher (*Jour. Clin. Inv., 23,* 365, 1944), Cournand, Noble, Breed, Lauson, Baldwin, Pinchot, and Richards (*Jour. Clin. Inv., 23,* 491, 1944), and Warren, Stead, Merrill, and Brannon (*Jour. Clin. Inv., 23,* 506, 1944).

F. Fibrinogen and Thrombin

The proteins concerned in the blood coagulation mechanism have been discussed by a number of authors. Recent reviews are given by Eagle (*Medicine, 16,* 95, 1937), Ham and Curtis (*Medicine, 17,* 413, 1938), Wöhlisch (*Ergebn. d. Physiol., 43,* 174, 1940), Ferguson (*Ann. Rev. Physiol., 2,* 71, 1940), Taylor, Davidson, and Minot (Nelson new loose leaf medicine, Vol. IV, Thomas Nelson & Sons, New York), Seegers (*Jour. Biol. Chem., 136,* 103, 1940), Seegers and McGinty (*Jour. Biol. Chem., 146,* 511, 1942), Neurath, Dees, and Fox (*Jour. Urol., 49,* 497, 1943), and Edsall, Ferry, and Armstrong (*Jour. Clin. Inv., 23,* 557, 1944). The purification and properties of prothrombin have been studied particularly by Mellanby (*Proc. Roy. Soc.,* London, *B 107,* 271, 1930) and by Seegers (*Jour. Biol. Chem., 136,* 103, 1940). Recent studies on the purification of thrombin have been carried out by Seegers and McGinty (*Jour. Biol. Chem., 146,* 511, 1942) and by Milstone (*Jour. Gen. Physiol., 25,* 679, 1942).

The fundamental studies on fibrin clots and fibrin films have been carried out by Ferry and Morrison (*Jour. Clin. Inv., 23,* 566, 1944), their use in the surface treatment of burns by Hawn, Bering, Bailey, and Armstrong (*Jour. Clin. Inv., 23,* 580, 1944), in the repair of dural defects by Bailey and Ingraham (*Jour. Clin. Inv., 23,* 597, 1944) and by Ingraham and Bailey (*Jour. Am. Med. Ass., 126,* 680, 1944). Fibrin foam was developed as a hemostatic agent by Bering (*Jour. Clin. Inv., 23,* 586, 1944). Its use in neurosurgery has been described by Ingraham and Bailey (*Jour. Neurosurg., 1,* 23, 1944), by Bailey and Ingraham (*Jour. Clin. Inv., 23,* 591, 1944), and by Woodhall (*Jour. Am. Med. Ass., 126,* 419, 1944).

G. Blood Grouping Globulins

The classical work on blood grouping is by Landsteiner (*Centralbl. f. Bakt., 28,* 357, 1900) (*Wien. Klin. Wochenschr., 14,* 1132, 1901). Blood grouping technique has recently been considered in detail by Schiff and Boyd (Blood grouping technic, Interscience Publishers, New York, 1942) and blood groups and transfusions by Wiener (Blood groups and transfusion, Charles C. Thomas, Springfield, Ill., 3d ed., 1943). For the separation and concentration of isohemagglutinins from group specific human serum or plasma see Thalhimer and Myron (*Jour. Am. Med. Ass., 118,* 370, 1942) and Pillemer (*Science, 97,* 75, 1943) and Pillemer, Oncley, Melin, Elliott, and Hutchinson (*Jour. Clin. Inv., 23,* 550, 1944), and for their appraisal see DeGowin (*Jour. Clin. Inv., 23,* 554, 1944).

H. γ-Globulin Antibodies (Immune Globulins)

The earlier literature on "The chemistry of antigens and anti-bodies" is reviewed by Marrack (*Medical Research Council,* Special Report Ser., No. 230, His Majesty's Stationery Office, London, 1938). The most recent investigations upon the chemical properties of the γ-globulins are considered by von Mutzenbecher (*Biochem. Zeitschr., 266,* 226, 250, 259, 1933), McFarlane (*Biochem. Jour., 29,* 407, 660, 1175, 1209, 1935), Kekwick (*Biochem. Jour., 32,* 552, 1938), Kabat and Pedersen (*Science, 87,* 372, 1938), Heidelberger (*Cold Spring Harbor Symp. Quat. Biol., 6,* 369, 1938; *Chem. Rev., 24,* 323, 1939), Kabat (*Jour. Exper. Med., 69,* 103, 1939), Cohn, McMeekin, Oncley, Newell, and Hughes (*Jour. Am. Chem. Soc., 62,* 3386, 1940), Cohn, Luetscher, Oncley, Armstrong, and Davis (*Jour. Am. Chem. Soc., 62,* 3396, 1940), Cohn (*Chem. Rev., 28,* 395, 1941; *Proc. Am. Philos. Soc., 88,* 159, 1944), Cohn, Oncley, Strong, Hughes, and Armstrong (*Jour. Clin. Inv., 23,* 417, 1944), Williams, Petermann, Colovos, Goodloe, Oncley, and Armstrong (*Jour. Clin. Inv., 23,* 433, 1944), the concentration of certain antibodies in globulin fractions derived from human blood plasma by Enders (*Jour. Clin. Inv., 23,* 510, 1944) and the use of the γ-globulin antibodies in the prophylaxis and treatment of measles by Stokes, Maris, and Gellis (*Jour. Clin. Inv., 23,* 531, 1944), by Ordman, Jennings, and Janeway (*Jour. Clin. Inv., 23,* 541, 1944), by Janeway (*Jour. Am. Med. Ass., 126,* 674, 1944), and by Greenberg, Frant, and Rutstein (*Jour. Am. Med. Ass., 126,* 944, 1944).

INDEX

Acetylcholine, 100–119
Achromotricia factor, 267
Adenosinetriphosphate, 115
Adiabatic fluid, 140, 141, 147
Adrenalin(e), 101, 109, 115, 118
"Adrenergic" fibers, 102
Adrian, 59
Agglutinogens, 296
Agranulocytosis, in rats, 271
Aileron pressure, 25
Air sickness, 40, 44
Albumin(s), 282–291; serum, 281, 289, 290
Almquist, 248
Altitude(s), high, 30–40; oxygen supply, 67–69
Ambiguity and groups, theory, 134, 135
American Coll., Physicians, 304
American Red Cross, 285, 303
Ammonium-iron-alum, 158; cell, 160; entropy, 163–166
Amytal, 193
Anemia, in dogs and pigs, 252
Anoxia, 32, 34–36
Ansbacher, 267
Anti-esterasics, 113
Aqueous humor, 84
Arago, 79
Archimedes, 123
Armstrong, 35
Arnold, 20
Ashburn 262, 271
Atom(s), dipole moments, 195; reactions, 169–176; energy, 147, 187; paramagnetic, 160; structure, 195
Atomic beams, 195
Atomic nucleus, 195
Atomic spin, 197
Atropine, 118
Aviation, high altitude, 30–40; instruments, 73; psychology, 1–48; World War, 1; sensory reactions, 73; medical service, 5, 20, 42
Aviation training, accidents, 21, 22; per cent elimination, 16; instruction, 23, 29; integration, 29; military, 21–28; plane, 23, 28; psychoörganic, 48; value, 28–30
Aviator(s), characteristics, 47; efficiency, 42–46; fighters, 6; forces affecting, 72; ideal, 47; indoctrination, 32; morale, 46; response habits, 40; selection, 6–21

Bacteria, 260, 261
Bancroft, 248
Barbital, 193
Barcroft, 108
Barnes, 77, 96
Barron, 108
Batschinski, 183
Baxter, 248
Beams, atomic, 195; molecular, 195; spectral frequencies, 199
"Bends," 36. *See also* Altitude(s)
Benzamide, 193
Bergson, 126
Bernard, 275
Bert, 66
Best, 254
Biotin, 260, 261–264; concentrate, 262; effect on Hamster, 263, 264; formula, 264
"Black care," 41, 43
Black tongue, 253
Blood, antibodies, 304; capillaries 276; characteristics, 286; clotting, 293–296; coagulation, 302–303; derivatives, 273–304; enzymes, 291–293, 302; functions and uses, 300–304; h(a)emoglobin, 30, 275, 277, 281, 301; hormones, 291–293, 302; lipids, 291, 302; products, 288; proteins, 291–293; public health record, 304; respiratory function, 274–278; transfusion(s), 290, 296, 303; types, 296, 297, 303
Blood plasma, 276. *See also* Plasma
Blood vessels, 281
Blood volume, control, 290–291, 301, 302; osmotic eeffct, 290
Body(ies), rigid, 121; elastic, 139; economy, 292; enzymes, 292; vitamin changes, 245–248
Bohr, 143; magneton, 168; relation, 199
Boll, 88
Bolometer, 75
Boltzmann 158, 179; principle, 177, 180, 187
Boole, 124
Bridgman, 131

Brimhall, 23
Brink, 61
British R.A.F., 41
Brown, G. L., 115
Brown, D. E., 189
Brumberg, 96
Buelbring, 110, 111, 112
Burch, 208
Burn, 110, 111, 112
Burrows, 248

"Caissons disease," 36. *See also* Altitude(s)
Calculus, absolute differential, 142
Calibration, retina, 93
Cannon, 101, 109, 110
Carotene(s), 245–248
Casimir, 162
Cawley, 247
Celanthrene red, 235, 243
Centrifugal still, 216–218; rotor, 219
Cerebrosides, 292
Chain of association, 120
Chariton, 77
Chase, 88, 89
Chemistry, theoretical, 169; vacuum, 205–248
Chloral hydrate, 193
Chloretone, 193
Cholesterol, 241
Cholinacetylase, 115
Choline, 100, 254–256; formula, 254
"Cholinergic" fibers, 102
Cholinesterase, 100, 101, 105, 110, 292
Chromotrichial factor, 262
Circle, squaring, 123
Civil Aeronautics Administration, 5, 23, 25, 28; "patter books," 24, 25
Claisen flask still, 218
Clotting, fibrin, 293; surgical, 293. *See also* Coagulation
Coagulation, 302, 303
Cocarboxylase, 250
Cole, 55
Colorimeter, 228
Columbus, Realdo, 273
Cornea, 84
Cotton, 156
Coulomb, law, 196, 203
Coulombic bonds, 186
Covalence, 170
Coxwell, 66
Crocé-Spinelli, 66
Crossley, 257
Cryptoxanthine, 246
Curare, 105, 106, 118
Curie, 152, 153; law, 165
Curiepoint, 164
Cypridina, 188
Cytochrome reductase, 251
Czerny, 77, 96

Daft, 262, 271
Dale, 101, 102, 103, 105
Dam, 248
Dartnall, 90
Davies, 62
Day, 271
DeBroglie, reflection, 195
Decomposition, molecular, 206, 209
Decompression chamber, 32, 33, 37
Deflection, magnetic moment, 197
Degassing vacuum, 211
de Haas, 156, 157, *162*
de Klerk, 162
de No, Lorentz, 59
Denis, 285
Densitometer, 220
Dermatitis, 252, 253, 263, 264
Descartes, 122, 124
Deuterium (D_2), 202, 203
Deuteron, 202; magnetic moment, 203; quadrupole electric moment, 202, 203
Dewar containers, 165, 166
Dewey, John, 120
Diamagnetic substances, 152
Diamylaminoanthraquinone, 235
Dicken, 270
Dickshit, 108, 111
Dimick, 255
Dipole moments, 195
Dirac, 147
Distilland, 209, 217; rate of travel, 221
Distillation, centrifugal, 215–219; cyclic, 231–234; high-vacuum, 248; molecular, 206; physics, 210; pilot dyes, 235; quantitative, 232; rate, 234; short-path, 207–208; vitamin A, 233–237, 246; vitamin D, 237–240. *See also* Oil and Still
Dixon, 99
DuBois, 30
Dudley, 101
du Vigneaud, 254, 262, 264
Dyes, distillation, 236, 243

E. coli. See Escherichia
Earth, isotropic sphere, 130
Eccles, 106
Effectors, nerve relations, 102, 116

Einstein, 76, 138, 140–142, 147; theory of gravitation, 147; theory of relativity, 138, 141
Electromagnetic space-time, 141
Electron, configuration, 171; pair, 170
Elliott, 99
Elvehjem, 253
Embree, 244, 245
Endothermal reactions, 185
Energy, quanta, 143; rate constants, 179; vision, 75
Engelhart, 102
Entropy, 158, 159, 163, 166, 193; melting, 183
Equilibrium, 169–194; blood, 279; limitation, 180; pressures, 175; reaction, 175; theory, 178
Equilibrium constant, 178
Ergotamine, 118
Ergotoxine, 118
Escherichia coli, 260, 261
Eserine, 101, 112
Esterase, 108
Estermann, 197
Ether, concept, 146
Euclid, 122, 123
Euclidean geometry, 124, 138
Evaporation, radiation losses, 224–227. *See also* Distillation
Evaporator, rotating, 224
Evipal, 193
Exothermal reactions, 185
Expansion, adiabatic, 149
Eye, dark adaptation, 78; diagram, 84; erg requirements, 75; media, 84; retina, 78; sensibility distribution, 87, 88; sensitivity, 78, 97
Eyster, 77, 96
Eyring, 180, 185, 187, 189, 193, 194

Fabricius, 273
Falling-film still, 213–218
Fallopio, 273
Faraday, 152
Fatigue, chronic, 44
Fawcett, 248
Feldberg, 115
Fibrin, clotting, 293; films, 294, 295; tubes, 295
Fibrin foam, 295; dry state, 295
Fibrinogen, 288, 295; action, 283; dimensions, 280, 281; mechanical properties, 293; plastics, 295
Fibrinolytic enzyme, 293, 296
Fitzgerald, 137
Flanagan, 15
Flannigan, 5
Flash still, 218
Flight recorder, 28
Fluid, adiabatic, 140, 141, 147
Flying, human, 2–6
Flying fatigue, 40–41
Flying Fortress, 45; crew, 29; fighting stations, 31; instrument panels, 27
Folch, 270
Folic acid, 260, 262
Folkers, 252
Formula(e), biotin, 264; choline, 254; niacin, 252; niacin amide, 252; pantothenic acid, 253; para-aminobenzoic acid, 267; phase-space, 188; pyridoxine hydrochloride, 253; riboflavin, 251; specific heat, 183; succinyl sulfathiazole, 259; sulfaguanidine, 257; vitamin A, 227, 242; vitamin A_2, 242; vitamin D, 239
Fourier, 130
Fovea centralis, 84
Fowler, 181
Frant, 300
Frauenhofer lines, 200
Frege, 122
Frisch, S., 96
Frisch, 197
Frost, 255

Gadolinium sulfate, 153
Gasser, 59
Galvanometer, 81
Gant, 260
Gas, adiabatic expansion, 160; kinetic theory, 205
Gas-liquefaction, 150; adiabatical, 151
Gellis, 300
Genetic foundation, 120
Geometry, 122; Euclidean, 124, 138; non-Euclidean, 122; postulates, 123; Riemannian, 142
Goethe, *Faust*, 274
Giauque, 151, 152, 156
Gillespie, 267
Glaisher, 66, 67
Globin, 277
Globulins, 288; alpha (α), 291; beta (β), 291; gamma (γ), 281, 283
Glucose, 115
Goldberger, diet, 271
Goodeve, 90
Gorter, 166

Granulocytopenia, 271
Green, 256
Greenberg, 300
"Gremlins," 41, 43
Guggenheim, 181
György, 262

H(a)emoglobin, 30, 275, 277, 281, 301
Haig, 88, 89
Haimovici, 109, 110
Hamilton's principle, 136, 143, 146
Hardy, 296
Harris, 247, 248, 252
Harvard University, 304
Harvey, E. N., 188
Harvey, William, 273, 275
Hauschildt, 265
Heart, donor, 99: recipient, 99
Hebb, 160
Hecker, 248
Heisenberg, 145
Helmholtz, 182
Hematocrit reading, 281
Hermite, 123
Heuristic conjecture, 148
Hirschfelder, 185
Hodgkin, 55
Hogan, 268
Homopolar bonds, 184, 185
Horace, *Odes,* 41, 43
Hunt, 100
Hydrogen, 201; heavy, 202; radiofrequency spectrum, 201
Hydrogen-deuterium (HD), 202

Inhibitor, luminescence, 190–194
Inositol, 264-267; anti-alopecia factor, 265
Ionic bonds, 184, 185
Iron, proteins, 277
Iron-ion, behavior, 157, 158, 160; energy, 162
Isbell, 270
Isoagglutinins, 287, 297, 303

Janeway, 300
Jenkins, 5
Jennings, 300
Johnson, 189, 193
Jukes, 253–255

Kamerlingh-Onnes, 152, 156
Kearns, 193
Keesom, 150, 151, 165
Kelly, 24
Kelvin, 130, 150, 167
Kepler, 144; planetary laws, 135
Kinetic, theory of heat, 131, 205
Kitol, 244, 245
Kitty Hawk, 4
Koenig, 86, 87, 91
Kries, 96
Kühne, 88
Kürti, 162

Lactobacillus casei, 259
Lagrangian equations, 136
Lainé, 162
LaMer, 194
Langevin, 152, 153
Langewiesche, 3
Langley, 75–77
Langmuir, 194
Langmuir-Knudsen equation, 230
Language, tensors, 142, 147; vectorial, 134
Laplace, 129
Larmor, 138, 142
Lattices, theory, 121
Laussac, 66
Lea, 77
Leeuwenhoek, 273
Leibniz, 123, 132, 134, 136
Lens, 84
Lepp, 255
Leucopenia, in rats, 271
Levi-Civita, 142
Lewis, 170
Li, 111
Liebig, 285
Ligamentum nuchae, 294
Light, color, 79; energy, 96; measurement, 75; transmission, 85
Lightnings, 45
Lindemann, 123
Logic, 124
Logical field, 122
London, F. W., 186
Lorentz, 137, 138; group, 138, 146
Lovelace, 38, 39
Luciferase, 188–194
Luciferase-luciferin, 188
Luciferin, 188–194
Ludvigh, 84, 85, 90
Luminescence, 188–190; inhibitors, 190–194
Lythgoe, 88, 89

MacIntosh, 107, 115
Magee, 189
Magnetic cooling, 159

Magnetic dipole, 195
Magnetic fields, 150–168, 199
Magnetic interactions, 203
Magnetic moment, 153, 154, 196, 197
Magnetism, rotational, 202
Magnets, electronic, 201
Mannering, 268
Manning, 266
Marcelin, 180, 188
Maris, 300
Marshall, 257
Marsland, 189
Martin, 262, 268, 270
Mashburn, 12; apparatus, 12; serial action, 13; test, 11
Mathematical theory, 124
Mathematics, four-vector language, 139, 141; physical theories, 120–151; pure, 122; quantum-mechanical, 131; three-vector language, 134
Maupertuis, 136
Maxwell, James Clerk, 126, 131, 136; law, 195, 197, 200
Maxwell Field, 13
McCarthy, 84, 85, 90
McCord, 127
McElroy, 193
McFarland, 34
McLeod gauge, 213
Measles, serum treatment, 300, 303
Mechanics, quantum, 145, 170, 176
Méchanique Céleste, 129
Melton, 16
Melville, 262
Mercury, 140, 141
Mesotrons, 203
Michelson, 137, 138, 146
Miles pursuitmeter, 9
Milieu intérieur, 275
Miller, 112
Minz, 111
Mitchell, 260, 270
Models, proximate, 220; ultimate, 130
Molecular beams, 195, 196–199
Molecular still, 208–214, 227, 245; research 229–234
Molecule(s), chain reaction, 173; collision, 173; decomposition, 206; deuterium, 202, dipole moments, 195; energy, 174, 176–178; fibrinogen, 293, 294; proteins, 275–304; reactant, 181; reflection, 195; resonance, 172; structure, 280–284; thermal motion, 206
Monoacetin, 193
Monomolecular oil layer, 211
Moore, 247
Morley, 137, 138, 146
Morse, 187
Muscarine, 103
Muscle, acetylcholine effect, 105

Nachmansohn, 104, 105, 107, 108, 115
National Research Council 5, 30, 304
Nerve(s), action, 106; augmentory, 117; inhibitors, 110, 117; optic, 84; reflexes, 109; stimulation, 98–119
Nerve cell(s), 49–72; content, 54, 55, 64, 65; cerebral, 64; organization, 61; in power-diving, 70; sensitivity, 55–58; sensory 50, 51, 70, 72
Nerve fiber(s), "adrenergic," 102; birefringence, 54; "cholinergic," 102
Nerve impulses, 55–58, 61; chemical transmission, 98–119. *See also* Acetylcholine, Cholinesterase, Eserine, and Adrenalin(e)
Nervous system, 50, 59, 98
Neurochemical mechanism, 117. *See also* Acetylcholine
Neutrons, 151, 195
Newton, 123, 131, 132, 134, 135, 145
Newtonian group, 139
Newtonian theory, 134, 141, 144, 145
Niacin formula, 252
Niacin amide, formula, 252
Nicotine, 103, 110, 118
Nicotinic acid, 252; effect on chick, 267
Nielsen, 260
Nordström, 140, 147; theory, 141
Norris, 265
Number, concept, 121
Nutrition, human, 269–272

Ocker, 73
Office of Scientific Research and Development, 304
Ohm, law, 182
Oil(s), absorption spectra, 243; distillation, 211, 219–224; dye distillation, 236, 243; fish-liver, 226, 227, 236, 240–245; "spiral stream," 221–223
Oil layer, distillation, 225
Optic nerve, 84
Ordman, 300
Organic anticipatory actions, 50
Organism, evolution, 51; effect of gravity, 71, 72
Oscillator, harmonic, 177

Osmotic pressure, 279, 290
Oxidase, 251
Oxygen, h(a)emoglobin function, 277, 278, 301; molecular behavior, 152; nerve cells, 63, 75; need, 30, 32; saturation, 30; starvation, 35
"Oxygen bottle," 32
Oxygen mask, 32, 69

Palmitate-stearate, 247
Pantothenic acid, 253–255; formula, 253
Para-aminobenzoic acid, 191, 193, 258, 267–269
Parachute jumping, experimental, 39
Parallelogram Law, 127–129
Paramagnetic atoms, 160
Paramagnetic salts, 153; saturation, 153
Pauli, 145
Pauling, 185
Pellagra, 253
Pelzer, 180
Phase-space formulation, 188
Phenobarbital, 193
Phosphatase, 292
Phosphatides, 292
Photobacterium phosphoreum, 189, 190
Photochemical equivalence law, 76, 92
Photoreceptors, 72
Physics, classical, 132, 135; nuclear, 195–204
Physical theories, mathematics, 120–149
Pigmentation, 253, 262
Pirenne, 78, 83, 96
Planck, 76, 142
Planck's constant, 180, 198, 199
Plankton, 245
Plasma, antibodies, 298–300; blood grouping, 296, 303; fractionation, 284–290, 295; pooled serum, 297, 304
Plasma proteins, 278–284, 301, 302; distribution, 287; drying, 285; fluid balance, 279, 301, 302; permeability, 284; storage, 302; structure, 280–284; value, 287
Plastics, compounds, 207
Poisson, 93–95; distribution curves, 95; probability distributions, 94
Polanyi, 180, 187
Polariographic method, 62
Polymer intermediates, 207
Pot still, 212, 217, 218
Poth, 260
Powell, 185
Pressure, aileron, 25; osmotic, 279, 290, 291
Pressure-density relations, 140
Principle of sufficient reason, 127, 131, 134, 148
Principle of uniformity, 131
Protein(s), blood, 273–304. *See also* Plasma
Prothrombin, 257, 258
Protons, 195
Purcell, 160
Pyelolithotomy, 293
Pyridoxine, 252–253
Pyridoxine hydrochloride, 252; formula, 253
Pyruvate, 115
Pythagoras, 120
Pythagorean theorem, 124

Quadrupole electric moment, 202, 203
Quanta of energy, 143
Quantum mechanics, 143, 145, 170, 176
Quantum theory, 76, 199
Quastel, 115
Quinizarin green, 235

Randolph Field, 13, 42
Rasmussen, 270
Rations, chick, 265—268. *See also* Dermatitis
Rations, synthetic, rats, 257, 259, 265
Reactions, atomic, 169–176; molecular, 173–194; chain, 173; heat, 185
"Reaction coördinate," 179
Reaction rates, 181–184; absolute, 194; prediction, 184–188; theory, 178–181
Red Cross Blood Donor Program, 298
Relativity theory, 138, 141
Retina, 78, 84, 85; energy calibration, 93; image, 92; scoptic luminosity 91; structure, 86
Retinal rods, 85–92; energy quanta, 91, 92
Riboflavin, 250–251; formula, 251
Ricci, 142
Richardson, 255
Riemann, 122, 142
Robertson, 65
Rockefeller Foundation, 304
Rod cell, 97
Roggenbau, 84
Rollin, 162
Rose, 262
Rotor still, 219; evaporator, 224

Rowntree, 21
Rubbo, 267
Russell, 122
Rutherford, 137
Rutstein, 300

Salicylamide, 193
Saslaw, 271
Schmitt, 53
Schrödinger, "wave equation," 143–145, 147
Schwab, 271
Schwann, 107
Sebrell, 262, 271
Seeing, frequency, 95–96; relative intensities, 95
Serum, 288; antibodies, 299; gamma (γ) globulin, 281; measles treatment, 300, 303; pooled, 297
Serum, albumin, 289; dimensions, 281
Shaffer, 19
Shantz, 244, 245
Sherrington, 107
Shigella paradysenteriae, 271
Shlaer, 78, 83
Shock, 279
Sieve, 267
Simon, 156, 162, 164, 166
Sivel, 67
Snell, 251, 252, 260
Space, absolute, 132, 133, 145; quantization, 195
Space-time, "astro-electric," 36, 146; curved, 142; electromagnetic, 141, 142; four-dimensional continuum, 143, 145; framework, 135, 138; three-dimensional, 134
Spectograph, ultra-violet, 244
Spectroscopy, 80, 176; radiofrequency, 197, 199–204
Spectrum, absorption, 88, 90; sensibility, 87, 88–92
Spicer, 271
Spin theory, 186
Square, group operations, 133; rotation operations, 133
Stavraky, 109, 110, 112
Stearn, 194
Stern, 195, 197
Steinhart 194
Still(s), behavior of liquids, 215; centrifugal, 226; centrifugal rim, 217; Claisen flask, 218; cyclic batch, 231–233; falling-film 213–218; flash, 218; glass pot, 212; 217, 218; molecular, 208–214, 227, 229–234, 245. *See also* Oil
Stiller, 253
Stokes, 300
Stokstad, 266
Strong, 251
Succinyl sulfathiazole, 259
Sulfaguanidine, 257
Sulfanilamide, 191, 192
Sun, 141
Synapses, postganglionic, 103
Synaptic transmission, 106, 113, 114

Temperature(s), absolute, 150; gas-liquefaction, 150; sub-zero, 36. *See also* Altitude(s)
Test(s), arrangements, 14; apperceptive, 18; rating, 16; civilian pilot, 18; score, 16; coördination, 7–13; discrimination, 19; Mashburn, 11; non-verbal, 15; prediction, 6, 17; preselection, 17; research, 18; psychomotor, 11, 14, 16; reaction time, 9, 10; School of Aviation Medicine, 14; selection, 6; Spearman's G, 15; Terman's Binet, 15; Thorndike's CAVD, 15; Thurstone's, 15; visualmotor, 10
Tetanus, 105
Theiler, 256
Theory, ambiguity and groups, 134, 135; Einstein, 147; equilibrium, 178; kinetic, 131, 205; lattices, 121; Newton, 141, 144, 145; Nordström, 141; quantum, 76, 199; quantum-mechanical, 143; reaction rate, 178; spin, 186; thermodynamic, 175
Thermal decomposition, 206
Thermal exposures, 218
Thermal motion, 152, 206
Thermodynamic theory, 175
Thermopile, 81
Thiamine, 250
Thiamine hydrochloride, 250
Thiamine pyrophosphate, 250
Thrombin, 288, 295, 293–296
Thunderbolts, 4
Time, absolute, 132, 133; epoch shift, 133; framework, 135, 138; invariable properties, 133; one-dimensional, 134. *See also* Space-time
Tissandier, 67
Tocopherol, 248
Topology, 149
Transmission, chemical, 98–119; inhibition, 110; synaptic, 106, 113, 114

Transmitter, liberation, 116
Trional, 193

United States Army Air Forces Aero Medical Laboratories, 38
Unna, 255

Vacuum, chemistry, 205–248; degassing, 211; nuclear physics, 195–204
Vagusstoff, 100, 101
van der Waals bonds, 184, 185
van Veen, 41, 43
van Vleck, 160
Variables, Hamilton, 136; quantitative form, 130
Variational principles, 147; topological, 148
Vavilov, 96
Velocity distribution, 197
Viljoen, 256
Vision, day, 78, 79; energy quanta, 75–97; night, 78, 79; owl, 78; stimulus, 79, 86; threshold, 75–77, 96
Visual purple, 88–92
Vitamin A, 240–248; absorption curve, 229; anhydro sub-, 229; bioassay, 230; blends, 228; distillation, 228, 229, 233–237, 246; esters, 229; formula, 227, 242; industry, 227–229; iso-anhydro, 229; pseudo, 229; storage, 226; sub-, 229
Vitamin A alcohol, 241
Vitamin A cholesterol, 241
Vitamin A_2, 241; formula 242
Vitamin B, complex, 249–272; "filtrate factors," 253; synthesis, 256, 257
Vitamin D, 237–240; formula, 239; redistillation, 238; structure, 239
Vitamin D_3, 241
Vitamin D_3 cholesterol, 241
Vitamin E, 241
Vitamin K, 256, 258
Vitamin changes, 245–248
Vitamin synthesis, 256–259, 261
Viteles, 5, 28
Vitreous humor, 84
von Kries, 77

Waisman, 253, 270
Wald, 87–89, 91
Waller, 163
Waterman, 248
Weaver, 88
Welch, 259
Welder, thermocouple, 225
Werner, coördination compound, 170
Wetthauer, 84
Whitehead, 122
Wiersma, 156, 157
Wigner, 180
Williams, R. E., 88
Williams, R. J., 253, 259, 260
Wilson, 271
Woolley, 253, 255, 265, 268, 270
Woolpert, 271
Woonton, 112
World, macroscopic, 174
Wright, L. D., 252
Wright Field, 38
Wynne-Jones, 180

Xanthine oxidase, 251
X rays, diffracted, 53

Zeeman effect, 197